Towards a deeper understanding of believing and achieving in educational settings:

Reciprocity and calibration of self-efficacy and academic performance

Kate Louise Talsma

BSci(Hons)

Submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy (Psychology)

in the Faculty of Health, School of Medicine

University of Tasmania, July, 2017

Declaration of Originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Signed:

Date: 25 July 2017

Authority of Access Statement

This thesis may be made available for loan and limited copying and communication in accordance with the Copyright Act 1968.

Signed:

Date: 25 July 2017

Statement of Ethical Conduct

The research associated with this thesis abides by the International and Australian codes on human and animal experimentation, the guidelines by the Australian Government's Office of the Gene Technology Regulator and the rulings of the Safety, Ethics and Institutional Biosafety Committees of the University.

Signed:

Date: 25 July 2017

Statement of Co-Authorship

The following people and institutions contributed to the publication of work undertaken as part of this thesis:

Author details:

Candidate: **Ms Kate Talsma**, School of Medicine (Psychology), University of Tasmania

Author 1: **Dr Kimberley Norris**, School of Medicine (Psychology), University of Tasmania, Australia

Author 2: **Assoc Prof Benjamin Schüz**, School of Medicine (Psychology), University of Tasmania, Australia

Author 3: **Prof Ralf Schwarzer**, School of Psychology, Freie Universität Berlin, Germany

Author roles:

Paper 1, **I believe, therefore I achieve (and vice versa): A meta-analytic cross-lagged panel analysis of self-efficacy and academic performance**

Located in Chapter 2

Candidate was the primary author (70%), and with authors 2 and 3 contributed to the idea, its formalisation and development, refinement and presentation (25%). Author 4 contributed to refinement and presentation (5%).

Paper 2, **Reciprocity in the self-efficacy ↔ academic performance relationship: The effect of measurement order**

Located in Chapter 3

Candidate was the primary author (70%), and with authors 2 and 3 contributed to the idea, its formalisation and development, refinement and presentation (30%).

Paper 3, **Miscalibration of self-efficacy and academic performance: self-efficacy ≠ self-fulfilling prophecy**

Located in Chapter 4

Candidate was the primary author (70%), and with authors 2 and 3 contributed to the idea, its formalisation and development, refinement and presentation (30%).

Paper 4, **When belief exceeds capacity: which factors differentiate over-efficacious students from their peers?**

Located in Chapter 5

Candidate was the primary author (70%), and with authors 2 and 3 contributed to the idea, its formalisation and development, refinement and presentation (30%).

We the undersigned agree with the above stated "proportion of work undertaken" for each of the above published (or submitted) peer-reviewed manuscripts contributing to this thesis:

Signed: ______________________

Date: 17th November 2017

Dr Kimberley Norris

Primary Supervisor

School of Medicine

University of Tasmania

Signed: ______________________

Date: 17th November 2017

Prof Ben Canny

Head of School

School of Medicine

University of Tasmania

Acknowledgements

This thesis only exists thanks to the contributions and support of a number of amazing people.

First and foremost is my supervisory team. I am deeply grateful for the guidance you have provided and I know I am a better researcher for it. Kim, not only have you brought your wealth of expertise and skills to my PhD experience, your support and enthusiasm have been central to its final completion. I am so grateful for your patience and direction in navigating through the PhD landscape. Ben, how can I shower you with verbose and flowery thanks without you wondering whether I have learned anything in the past few years? Thank you for bringing your considerable research and statistical know-how to this project, and for providing valuable feedback throughout my candidature; this has contributed to a rigorous research training experience.

Thanks, Ralf, for your valuable input into the original meta-analysis – and for making me feel like getting papers rejected means I'm a member of some cool club rather than a failure.

Many thanks to the authors of the papers included in the meta-analyses in this thesis; without your interest in this project, and willingness to assist with the provision of additional information, the analyses would not have been possible.

To the 2014 and 2015 cohorts of first-year psychology students at UTAS, thank you so much for volunteering your time to participate in the longitudinal surveys which provided the data for the primary studies of this thesis.

My appreciation also goes to the University of Tasmania and the Australian Government, for their financial support during my candidature. Thank you to the Divisional staff who helped make this possible, and to the cohort of students who shared this PhD journey.

Thanks also to the beautiful friends and family who have supported me these past few years, providing both consolations and congratulations at key moments.

Aiden, my precious boy, thank you so much for your love, hugs and patience during this process. Your creativity, exuberant character, and growing independence are wondrous to me. Thank you for keeping me grounded, for the many laughs we share, and for the constant reality checks you provide.

Michael, I cannot thank you enough for your love and support during this journey. Thank you for the innumerable cups of tea (and more than a few glasses of wine), for reading my drafts so good-naturedly, and for the occasional playful reminder not to take everything too seriously. You are my rock. Your unwavering belief in me – coupled with your cheeky humour – have kept me together and I would not have finished this thing without you. I'm looking forward to sharing whatever life has in store for us next.

Abstract

Background: Academic self-efficacy, and its relationship with academic outcomes, has been the subject of extensive research and many systematic reviews over the past several decades. These analyses consistently point to self-efficacy as one of the strongest positive correlates of academic performance, reflecting the traditional focus on the benefits of a strong sense of self-efficacy on educational behaviours and performance. However, questions about the nature of the relationship between self-efficacy and performance remain, including the issue of reciprocity in the relationship and the relative impact of performance on subsequent self-efficacy beliefs, and the issue of how accurate people's self-efficacy beliefs are and how this influences performance.

Research questions: In this context, the overall aim of the present thesis is to explore the complexities in the relationship between self-efficacy and academic performance, moving beyond the notion of "believe and you will achieve". To this end, this thesis addresses the following research questions. Firstly, what is the impact of academic performance on self-efficacy and vice versa: is the relationship between the two reciprocal, or is one variable a stronger antecedent of the other? And is the relationship moderated by other participant and methodological variables? Secondly, how accurate are students' self-efficacy beliefs and how does this accuracy (or inaccuracy) predict future academic performance outcomes? And is it possible to identify person and environmental characteristics which differentiate miscalibrated students from their peers?

Studies: Studies 1 and 2 addressed the question of reciprocity in the self-efficacy/ performance relationship by means of a systematic review with meta-analysis of panel data. Pooled correlations were fit to a cross-lagged path model which provided support for reciprocal effects. Moderator analyses indicated that the strength of cross-lagged effects varied as a function of participant age, the lag time between measurements, the degree of

match between self-efficacy and performance operationalisations, and the type of scale used to measure self-efficacy. Overall, reciprocity held in most circumstances, providing support for reciprocal determinism as per self-efficacy theory, and highlighting the positive influence of self-efficacy on academic performance and vice versa. A key finding was that reciprocity was evident in adult samples, but not in children (in whom performance predicted self-efficacy, but not vice versa). Also, the order of measurement of the variables at each measurement point moderated the reciprocal relationship. When performance was measured prior to self-efficacy at time 1, self-efficacy was a stronger predictor of performance at time 2 than the reverse. When self-efficacy was measured first at time 1, performance was a stronger predictor of self-efficacy at time 2 than the reverse.

Study 3 focused on calibration of self-efficacy and academic performance in university students, operationalised as the deviation of self-efficacy beliefs from performance outcomes measured on the same scale. Participants' self-efficacy beliefs with regard to their performance could be accurate (calibrated), or inaccurate (miscalibrated), and miscalibration was further categorised as either over- or under-efficaciousness. Miscalibration was prevalent, with under-efficaciousness evident at task-level (written assignments and exams) and over-efficaciousness pronounced at subject-level (overall grades). Low achievers tended to be over-efficacious, while the reverse was true for high achievers. The strongest subsequent performance outcomes on similar tasks were predicted by under-efficaciousness, rather than accuracy or over-efficaciousness. These findings suggest that over-efficacious students may be at risk of negative academic outcomes. Findings were more consistent with discrepancy-reduction processes than with the idea of self-efficacy as a self-fulfilling prophecy.

Study 4 investigated whether it is possible to identify over-efficacious students (identified in study 3 as being at risk of poor academic performance) based on personal and

environmental variables. Self-efficacy calibration was assessed at multiple time points over an academic semester. At different points in the semester, over-efficacious students tended to be younger, lower in cognitive ability and agreeabless, higher in self-esteem, and from higher SES backgrounds. Over-efficaciousness may, thus, be maintained by both cognitive and motivated processes. These findings may be of benefit in targeting interventions to enhance students' calibration levels, and they may also inform future research and theory development.

Conclusions: Overall, the findings of this thesis present a picture of the relationship between self-efficacy and academic performance as complex and nuanced. Reciprocity between the two variables is evident, but varies in magnitude and directional strength with different participants and research approaches. Calibration in adult learners appears to be poor overall, and the direction of miscalibration predicts future performance. Both personal and contextual variables differentiate over-efficacious students from their peers. Future research and theory development, and applications of self-efficacy theory to learning environments, would benefit from increased focus on the issues of reciprocity and calibration, which point to the need for updated conceptualisations of adult learners' self-efficacy beliefs, as well as to the importance of regular cycles of performance and accurate feedback for students of all ages.

Table of contents

List of Tables

List of Figures

Chapter 2

Chapter 1: Introduction

Self-efficacy is one of the most extensively researched constructs in psychology. Bandura's (1977) seminal article introducing the concept of self-efficacy has been cited in excess of 50,000 times, with research exploring the construct in different age groups and cultures, in relation to different domains and tasks, and in settings ranging from workplaces and classrooms to hospitals and sports fields (Schunk & Pajares, 2004). Self-efficacy refers to "people's judgements of their capabilities to organise and execute courses of action required to attain designated types of performance outcomes" (Bandura, 1986, p. 391).

Self-efficacy sits at the core of social cognitive theory, which posits that human functioning is a dynamic interchange between person factors (including self-efficacy) along with behavioural and environmental factors, which are hypothesised to influence each other reciprocally (Bandura, 1986). Thus, convictions of ability to perform are believed to be associated with a range of behavioural factors, such as the choice of activities one will pursue, the amount of effort likely to be expended, persistence in the face of challenges, and, ultimately, the performance outcome itself (Bandura, 1997). Sources of information which are posited to form the basis of self-efficacy judgements include *mastery experience*, based on actual performance success or otherwise, *vicarious experience*, that is, the observed performance of others, *verbal or social persuasion*, including encouragement from important others, and *physiological states*, such as when intense arousal in advance of performing on a task might be interpreted as being caused by a lack of ability to perform (Bandura, 1997). Mastery experience is considered to be the strongest source of these four, with successful performance resulting in increased self-efficacy and failure decreasing it (Schunk & Pajares, 2004).

There are several constructs which overlap with self-efficacy, but which at the same time can be differentiated from it. This is important, as terms are sometimes used interchangeably in the literature even though they refer to different constructs. Self-efficacy refers to a person's conviction that they have the ability to perform a behaviour in the future (Schwarzer & McAuley, 2016); it has been termed an *I can* construct (I can jump that puddle). This differentiates it from self-concept, as a past-oriented self-schema with an *I am* focus (I am good at puddle jumping) (Bong & Skaalvik, 2003). Self-efficacy can also be differentiated from self-esteem, which is a global evaluation of one's self-worth (Rosenberg, 1979). Self-efficacy is sometimes measured as a generalised, enduring, trait-like construct that refers to a person's belief in their broader coping and problem-solving abilities across multiple contexts (e.g., Judge, Erez, Bono, & Thoresen, 2002; Scholz, Doña, Sud, & Schwarzer, 2002). However, this is a different conceptualisation of self-efficacy than that espoused by classic self-efficacy theory (Bandura, 2012; Schwarzer & Jerusalem, 1995). In the traditional view, self-efficacy is considered to be domain-specific (e.g., mathematics self-efficacy) or task-specific (e.g., self-efficacy for specific mathematics problems such as solving equations). Typically, the more specific a measure of self-efficacy is, the more strongly it predicts related performance outcomes (Bong, 2001; Pajares & Miller, 1995).

Confidence is another construct that may be confused with self-efficacy. Confidence is defined as "a state of being certain about the success of a particular behavioural act" (Stankov, Lee, Luo, & Hogan, 2012, p. 747). Unlike self-efficacy, which is traditionally measured prior to the behaviour in question, confidence is typically measured concurrently with, or subsequently to, performance (Schraw, 2009; Stankov & Lee, 2008). As such, confidence refers to a belief about something that has just been done, while self-efficacy refers to a belief about something that is yet to be done. Also, although confidence ratings are generally made at an item level (i.e., participants respond to a test item and then indicate their confidence that their answer is correct as a percentage), the construct itself is considered to be a global trait which falls somewhere between ability and personality (Stankov, Kleitman, &

Jackson, 2015; Stankov & Lee, 2008; Stankov et al., 2012). This also differentiates it from self-efficacy, which is typically considered to be domain- or task-specific as outlined above.

Self-efficacy can also be distinguished from self-regulation, an important and widely-researched construct in educational psychology. Self-regulation is “the self-directive process by which learners transform their mental abilities into academic skills” (Zimmerman, 2002, p. 65). Self-efficacy is one of the beliefs which underlies an individual’s self-regulatory capacity. In Zimmerman’s three phase model of self-regulation, self-efficacy forms part of the pre-performance phase, where it contributes to a person’s motivation in combination with other variables such as outcome expectations and task value (Zimmerman, 2000, 2002). Self-regulation also includes a performance phase, in which the individual engages in self-monitoring and control processes, and a post-performance phase, in which the individual engages in attribution and reflection processes.

Self-efficacy is argued to “touch virtually every aspect of peoples’ lives” as a critical and potent influence on achievement (Pajares, 2006; p. 341). Early research on the self-efficacy construct focused on behaviour change in snake phobics, in which self-efficacy for performing a feared behaviour, such as touching a snake, was reported to be more predictive of behaviour than previous behaviours (e.g., Bandura, 1977). Following on from this early tradition, a key line of inquiry with regard to self-efficacy is its relationship with performance outcomes. For example, meta-analyses have shown that self-efficacy is positively correlated with job performance (Judge & Bono, 2001; Stajkovic & Luthans, 1998) performance in sport (Moritz, Feltz, Fahrbach, & Mack, 2000) and performance on memory tasks (Beaudoin & Desrichard, 2011).

Academic self-efficacy refers to individuals’ judgements about their capacity to perform in academic settings. Exploration of academic self-efficacy has been one of the most prolific research pursuits in this area (Klassen & Usher, 2010). Research exploring the relationship between academic self-efficacy and academic performance provides support for the predictions of self-efficacy theory outlined above. For example, research suggests that students with higher levels of self-efficacy set themselves more challenging academic goals, which leads in turn to better academic achievement (Brown et al., 2008; Zimmerman, Bandura, & Martinez-Pons, 1992). Students with higher levels of academic self-efficacy are also more likely indicate willingness to persevere with learning tasks in spite of boredom or difficulty, which also leads to stronger academic performance (Komarraju & Nadler, 2013).

Academic self-efficacy is reported to account for approximately a quarter of the variance in academic outcomes (Klassen & Usher, 2010; Pajares, 2006); given the complex nature of the learning process and the range of variables likely to be implicated, an effect of this size is considerable. The first of a string of meta-analyses conducted on the relationship between self-efficacy and academic performance was undertaken more than 25 years ago and is still influential today (Multon, Brown, & Lent, 1991). In this review, which pooled results from 36 studies, the authors reported a correlation between self-efficacy and academic performance of r=.38, accounting for approximately 14% of the variance in academic outcomes across a range of student ages and performance measures. Self-efficacy was also found to be the strongest non-intellective correlate of university GPA in two more recent meta-analytic reviews. Robbins et al. (2004) reported a spearman rank correlation of ρ=.496 based on 19 studies, and Richardson, Abraham, and Bond (2012) reported a sample-weighted average correlation of r=.59 based on 4 studies. Most recently, Honicke and Broadbent (2016) reported a mean sample-weighted correlation of r=.33 between academic self-efficacy and performance in university students, based on 51 studies. Thus, while there are many constructs which are relevant to understanding academic performance, self-efficacy has consistently been identified as one of the most influential.

Since the earliest reviews were conducted, findings regarding the positive relationship between self-efficacy and academic performance have underscored recommendations to conduct self-efficacy-enhancing interventions in learning settings, with the goal of improving academic performance (e.g., Multon et al., 1991; Richardson et al., 2012). Similar interpretations can be seen with regard to work performance (e.g., Judge & Bono, 2001; Stajkovic & Luthans, 1998). Whether this tendency was born in the research world or in applied environments themselves, the majority of self-efficacy research has focused on the potential benefits of a strong sense of self-efficacy on academic performance outcomes. Thus, while the majority of research in this area is cross-sectional, interpretations of findings tend to focus on self-efficacy as cause and performance as effect, and this approach has permeated the body of literature considering the relationship between the two variables. Self-efficacy is argued to be at the core of a self-fulfilling prophecy, where what you believe leads to what you achieve (Pajares, 2006). In educational settings, it is suggested that this self-fulfilling prophecy means that students with high self-efficacy persevere and perform well, while those with low self-efficacy give up and disengage (Pajares, 2002). This perspective has guided much of the research on the topic, such that, although a vast body of literature exists regarding self-efficacy and academic performance, there are complexities in the relationship yet to be explored. The overall theme of the present thesis is the exploration of these complexities in the relationship between self-efficacy and performance, focusing on dynamics which may not be able to be explained by the traditional “believe and you will achieve” perspective, in which academic self-efficacy is viewed as the antecedent of academic performance.

To this end, two lines of research which warrant further investigation are, firstly, reciprocity and the comparative strength of directional effects in the relationship between self-efficacy and academic performance, and secondly, the matter of the accuracy of self-efficacy beliefs; that is, the degree of concordance or calibration between self-efficacy and academic performance. Both of these issues are important because findings may suggest that widely held conceptualisations of self-efficacy, and its association with academic performance, oversimplify the nature of the relationship between the two. The issues of reciprocity and calibration are considered in more detail below, and the background and aims of the studies of the present thesis are outlined.

Reciprocity in the relationship between self-efficacy and performance and the strength of directional effects

The association between self-efficacy and academic performance may be explained by the influence of self-efficacy on academic performance, the influence of academic performance on self-efficacy, or a combination of the two in the form of reciprocity (Valentine, DuBois, & Cooper, 2004). According to self-efficacy theory, cognitions (including self-efficacy beliefs) and behaviours influence each other reciprocally (Bandura, 1978). Research into the relationship between self-efficacy and academic performance has tended to focus on unidirectional elements of the reciprocal relationship in isolation. As noted above, the majority of research considers the relationship from a self-efficacy → performance perspective (e.g., Brown et al., 2008; Caprara, Vecchione, Alessandri, Gerbino, & Barbaranelli, 2011a; Robbins et al., 2004). A smaller body of research considers the influence of performance on self-efficacy – generally from the perspective of mastery experience as a source of self-efficacy beliefs (e.g., Phan, 2012). Others focus specifically on an argument that the cross-sectional findings may be better explained by the influence of performance experience on self-efficacy than the reverse (Richard, Diefendorff, & Martin, 2006; Vancouver, Thompson, & Williams, 2001). In both cases, research in the performance → self-efficacy paradigm typically shows that academic performance outcomes positively predict self-efficacy beliefs (Klassen, 2004). Some researchers have gone so far as to suggest that self-efficacy constitutes a reflection of previous performance and is of little predictive value when considered independently (Heggestad & Kanfer, 2005). A small number of studies stagger measurements of self-efficacy and academic performance in order to investigate the issue of causal precedence in the relationship (e.g., Caprara, Vecchione, Alessandri, Gerbino, & Barbaranelli, 2011b). Some researchers focus on ascertaining the longitudinal effect of self-efficacy on subsequent performance while controlling for previous performance (Valentine et al., 2004).

However, in spite of the merit of these approaches, research in this area has not convincingly determined the relative influence of self-efficacy and academic performance on each other over time, nor has it shown unequivocally that the relationship between self-efficacy and academic performance is indeed reciprocal, as posited by self-efficacy theory. In addition, controversy persists regarding the best approach to assessing these effects (Bandura & Locke, 2003; Feltz, Chow, & Hepler, 2008; Heggestad & Kanfer, 2005; Vancouver et al., 2001). Given that important theoretical positions rest on a causal influence of self-efficacy on subsequent performance, and that this conceptualisation underlies widespread application of research findings in learning settings, it is critical to ascertain whether the weight of evidence supports this position (Valentine et al., 2004).

Based on the background outlined above, the aim of study 1 (chapter 2) was to explore reciprocity and the comparative strength of directional longitudinal effects in the relationship between self-efficacy and academic performance. Models which include simultaneous measures of both self-efficacy and academic performance over two measurement waves (i.e., panel designs: two-wave, two-variable designs) provide data which can be used in a more rigorous test of reciprocal relationships than has previously been conducted (Selig & Little, 2012). Thus, to address this chicken-and-egg question, I systematically searched the literature for panel studies measuring both self-efficacy and academic performance, and fit meta-analytic effect sizes to a cross-lagged panel model as illustrated in Figure 1, Chapter 2. Positive cross-lagged path coefficients reflect reciprocity, with the strongest path (if any) representing the strongest antecedent in the relationship. This approach enabled the determination of the unique influence of self-efficacy at time 1 on academic performance at time 2, and the unique influence of academic performance at time 1 on self-efficacy at time 2, holding other model paths constant. It was anticipated that the relationship between the two variables would be reciprocal, but because of research evidence showing positive effects for both self-efficacy → performance and performance → self-efficacy, the relative strength of these effects was approached in an exploratory fashion. Based on previous literature reporting a range of moderator effects (Honicke & Broadbent, 2016), I also investigated how the relationship varied as a function of sample characteristics including the age and sex of participants, and methodological characteristics including the concordance between the self-efficacy measure and performance task, the time elapsed between measurement waves, and the nature of the scale used to measure self-efficacy. Given the nature of the review conducted in chapter 2, this forms the primary literature review for the present thesis.

Chapter 2 as presented is currently under review as:

Talsma, K., Schüz, B., Schwarzer, R., & Norris, K. (2017). I believe, therefore I achieve (and vice versa): A meta-analytic cross-lagged panel analysis of self-efficacy and academic performance. *Learning and Individual Differences*.

Study 2 (chapter 3) followed the same basic procedure as study 1 (chapter 2), but considered the order of measurement of the two key variables at each measurement wave. In study 1, inclusion criteria specified that self-efficacy be measured prior to performance at each measurement wave, in line with recommendations from self-efficacy theory (Moriarty, 2014). However, previous research suggests that the strength of the relationship between self-efficacy and academic performance may differ depending on the order in which the variables are measured (Multon et al., 1991). This being the case, effect sizes in the posited reciprocal relationship between the two variables may also be affected by measurement order. The identification (during the literature search for study 1) of a number of studies in which measurement order was reversed provided the opportunity to explore measurement order as a moderator of the reciprocal relationship between self-efficacy and performance, and to examine potential differences in cross-lagged effects based on measurement order. Based on theoretical models of the relationship between self-efficacy and performance (Bandura, 1986; Gist & Mitchell, 1992) which suggest that self-efficacy beliefs are informed by analyses of performance experiences, it was anticipated that when performance was measured prior to self-efficacy at time 1, self-efficacy would be a stronger predictor of self-efficacy at time 2 than when self-efficacy was measured prior to performance at time 1.

Chapter 3 as presented is currently awaiting submission as:

Talsma, K., Schüz, B., & Norris, K. (2017). Reciprocity in the self-efficacy ↔ academic performance relationship: the effect of measurement order. *Manuscript in preparation*.[1]

---

[1] This article builds on Chapter 2 (study 1) and refers to it directly to orient readers to relevant conceptual and methodological issues, and compares model paths in the study 1 and study 2 articles. For this reason, this article (Chapter 3, study 2) has not yet been submitted for publication, as it is necessary for Chapter 2 (study 1) to be published first.

Together, studies 1 and 2 (chapters 2 and 3) aim to provide key insights into the dynamic nature of the relationship between self-efficacy and academic performance, and the situations under which the strength of bidirectional effects may vary.

Differences in the strength of directional associations in the self-efficacy ↔ academic performance relationship based on measurement order have been suggested to rest on differences in concordance between self-efficacy and performance depending on when the variables are measured (e.g., Multon et al., 1991). That is, self-efficacy may be more or less strongly related to previous or future performance (as per the cross-lagged paths in Figure 1, Chapter 2) depending on how calibrated the two variables are with each other. At the same time, stability of self-beliefs and academic performance over multiple measurements appears to be very strong (Foster, Was, Dunlosky, & Isaacson, 2016). These points raise a question as to the function of the feedback loop suggested to characterise the relationship between self-efficacy and performance (Gist & Mitchell, 1992; Lindsley, Brass, & Thomas, 1995), which leads us to the second key focus area for the present thesis.

Calibration of self-efficacy and academic performance

It is generally suggested that academically stronger students have higher self-efficacy beliefs, while academically weaker students have lower self-efficacy beliefs (Schunk & Pajares, 2004). However, this approach of comparing self-efficacy levels between students does not take into account how the self-efficacy beliefs of any given individual relate to their *own* performance capacity. This relates to *calibration* of self-efficacy, which is not concerned with whether individuals have high or low self-efficacy compared to others, but with whether an individual's self-efficacy beliefs over or underestimate their own performance capacity. To elaborate, the self-efficacy beliefs of high-achieving student Alex may be stronger than those of low-achieving student Andy, but they may simultaneously fall short of Alex's own capacity to perform, while in contrast, Andy may be over-efficacious.

Calibration of self-efficacy beliefs is important because of the role of self-efficacy in academic self-regulation and its influence on performance outcomes (e.g., Dunlosky & Rawson, 2012). An over-efficacious student may be overly ambitious in choosing challenges and meet with failure as a result, or may alternatively underestimate the amount of effort or preparation that is required to successfully perform a task (Boekaerts & Rozendaal, 2010; Schunk & Pajares, 2004). Most commentary on this issue relates to concerns about how students with illusorily positive views lack the necessary realistic foundation from which to approach their learning in educational settings (e.g., Moore & Healy, 2008; Valentine et al., 2004). However, under-efficaciousness may also be associated with inefficient academic self-regulation. An under-efficacious student may avoid attempting challenging tasks and thus prevent skill development, or may misallocate resources by over-studying in certain areas (Boekaerts & Rozendaal, 2010; Schunk & Pajares, 2004).

Calibration research has been conducted using a number of different combinations of measures which need to be differentiated conceptually. A considerable body of research has considered calibration of metacognitive judgements (see Schraw, 2009 for a review and taxonomy). Metacognitive judgements can include judgements made prior to performance (such as “judgements of learning” and “feeling of knowing”), during performance (sometimes called “online” accuracy or confidence judgements), and after performance (which can include retrospective accuracy or confidence judgements). Calibration research has also been conducted investigating understanding of texts, known as calibration of comprehension (e.g., Stolp & Zabrucky, 2009), as well as the accuracy of test performance predictions (e.g., Hacker, Bol, Horgan, & Rakow, 2000).

While a large body of research considers calibration, very little research specifically considers the calibration of beliefs which form part of a social-cognitive framework (Stolp & Zabrucky, 2009), such as self-efficacy. Given the importance of these self-beliefs to effective academic self-regulation (Zabrucky, 2010; Zimmerman, 2002), further research regarding self-efficacy calibration is critical. While a small number of studies have been conducted examining the accuracy of self-efficacy beliefs (Chen & Zimmerman, 2007; Chen, 2003), there are many gaps in the literature: for example, little is known about the calibration of university students self-efficacy beliefs, calibration of self-efficacy in domains other than mathematics, and for tasks other than tests. There is also a scarcity of research regarding self-efficacy calibration for academic performance tasks undertaken within naturalistic learning settings – that is, tasks assessed as part of a course of study as opposed to conducted in a laboratory setting or designed only for the purpose of an experiment (Hacker, Bol, & Keener, 2008). In the same vein, there is a dearth of research considering the impact of calibration on future performance outcomes (Bol & Hacker, 2012). In regard to this question, the optimal discrepancy between self-efficacy and capacity is a matter of considerable debate: some argue that accuracy of self-efficacy beliefs should best predict future academic performance (Stankov & Lee, 2017), while others argue that under-efficaciousness (Vancouver & Kendall, 2006) or over-efficaciousness (Pajares, 2006) are most likely to be the ideal recipe for academic success.

Study 3 (chapter 4) seeks to address gaps in the calibration literature by exploring the calibration of university students' self-efficacy beliefs with respect to authentic academic performance outcomes at task level (exams and written assignments) and at subject level (overall subject grades). I considered both self-efficacy *accuracy*, which reflects the absolute deviation between self-efficacy and an academic performance outcome measured on the same scale, and self-efficacy *bias*, which refers to the direction of that deviation in terms of either over- or under-efficaciousness. Study 3 considered the relative prevalence of calibration and miscalibration of self-efficacy across multiple academic tasks over the course of a semester. Based on previous findings of overconfidence in the broader calibration literature (e.g., Ehrlinger, Johnson, Banner, Dunning, & Kruger, 2008; Kruger & Dunning, 1999; Stankov & Lee, 2014; Stankov, Lee, & Paek, 2009; Zabrucky, 2010), it was anticipated that students would be over-efficacious overall across the academic outcomes measured.

An important contribution of chapter 4 is its response to recent calls in reviews of the calibration literature for the examination of the influence of calibration on one occasion on analogous performance outcomes on a subsequent occasion (Bol & Hacker, 2012). Using two types of authentic academic tasks, written assignments and exams, regression analysis was used to predict performance at time 2 from self-efficacy accuracy and self-efficacy bias at time 1. Given the debate about what level of self-efficacy is optimal for academic success, this question has a bearing on arguments for and against the use of programs in educational settings designed to enhance self-efficacy beliefs (Valentine et al., 2004). As all three potential hypotheses regarding the prediction of performance from calibration (i.e., accuracy *versus* over-efficaciousness *versus* under-efficaciousness) have a basis in previous research and theory, this question was approached in an exploratory manner.

Chapter 4 as presented is currently under review as:

Talsma, K., Schüz, B., & Norris, K. (2017). Miscalibration of self-efficacy and academic performance: Self-efficacy ≠ self-fulfilling prophecy. *Zeitschrift für Pädagogische Psychologie.*

Building on study 3 (chapter 4), study 4 (chapter 5) sought to explore which variables characterise those individuals whose self-efficacy beliefs diverge from their objective performance outcomes. Illusory self-beliefs have been identified as a risk factor for academic outcomes (e.g., Dunlosky & Rawson, 2012; Valentine et al., 2004), but little is known about what differentiates over-efficacious students from their peers. Logistic regression analysis enabled an analysis of whether a range of person variables (e.g., personality, self-beliefs) and contextual variables (e.g., SES, educational background) were associated with over-efficaciousness in adult learners with regard to subject-level outcomes, when considered at different points over an academic semester. Identification of variables which characterise over-efficacious students has implications for theory development and for targeted calibration interventions in learning settings. As there is minimal previous research in this area, and a lack of a theoretical model regarding self-efficacy calibration, backward elimination was used in regression models exploring this question. This study provides important preliminary insights into the individual differences that tend to accompany over-efficaciousness.

Chapter 5 as presented is currently under review as:

Talsma, K., Norris, K., & Schüz, B. (2017). When belief exceeds capacity: Which factors differentiate over-efficacious students from their peers? *Personality and Individual Differences.*

Chapter 6 summarises and integrates the findings of the studies which comprise the thesis, and discusses limitations, implications for research and practice, and directions for future research.

Reference lists which accompany articles under review (chapter 2/study1, chapter 4/study 3 and chapter 5/study 4) are located with the respective articles, while references associated with chapters not under review (chapter 1, chapter 3/study2, and chapter 6) are located at the end of the thesis.

Chapter 2: Study 1

I believe, therefore I achieve (and vice versa):

A meta-analytic cross-lagged panel analysis of self-efficacy and academic performance

Abstract

Self-efficacy has long been viewed as an important determinant of academic performance. A counter-position is that self-efficacy is merely a reflection of past performance. Research in the area is often limited by unidirectional designs which cannot address questions of reciprocity or the comparative strength of directional effects. This systematic review and meta-analysis, therefore, approaches the relationship as a chicken-and-egg question, considering both directions of the relationship simultaneously by pooling data from longitudinal studies that measured both academic self-efficacy and academic performance over two waves. Pooled correlations from 11 studies (N=2,688) were subjected to a cross-lagged path analysis that provided support for a reciprocal effects model. Performance had a net positive effect on subsequent self-efficacy (β=.205, *p*<.001), significantly larger than the effect of self-efficacy on performance (β=.071, *p*<.001). The unique impact of performance on subsequent self-efficacy aligns with the proposition that performance experiences influence self-efficacy judgements, whereas the unique impact of self-efficacy on subsequent performance supports a mobilisation effect. Moderator analyses indicate that the self-efficacy ↔ performance relationship varies as a function of participant age, the lag time between measurements, the degree of match between self-efficacy and performance operationalisations, and the type of scale used to measure self-efficacy. A key finding is that reciprocity holds for adults, but not for children: performance uniquely impacts subsequent self-efficacy beliefs, but not the reverse. Cross-lagged effects were stronger where studies used methodologies consistent with recommendations of self-efficacy theorists. Implications for theory, research, and practice are discussed.

Key words: *self-efficacy; academic performance; reciprocal effects; cross-lagged panel analysis; meta-analysis*

The non-intellective antecedents of student performance are of great interest to educators and education researchers (Robbins et al., 2004; Stankov & Lee, 2014), and research in this area is an important determinant of education policy (Bong, 2012; Pajares & Usher, 2008). One construct which has received a great deal of research attention is perceived self-efficacy – a core dimension of human agency widely believed to be positively related to academic success (Bandura, 1977, 1997; Bong, 2012; Pajares & Schunk, 2001). Self-efficacy refers to an individual's perception of their own capability to organise and execute required courses of action to achieve particular outcomes (Bandura, 1977, 1997). Self-efficacy is believed to enhance performance through a range of mechanisms: individuals with high levels of self-efficacy set more difficult goals, expend more effort, persist for longer with challenges, and show resilience in the face of adversity (Klassen & Usher, 2010). These achievements in turn are assumed to increase self-efficacy, which results in a self-fulfilling prophecy process (Bandura, 1977, 1997).

Self-efficacy → academic performance (I believe; therefore I achieve)

A vast body of research has explored the idea that self-efficacy is the antecedent in the relationship and exerts a positive motivational influence on performance (Honicke & Broadbent, 2016; Vancouver, Thompson, & Williams, 2001). Such research takes its lead from early studies by Bandura and colleagues (see Pajares, 1997; Zimmerman, 2000 for review), which demonstrated that self-efficacy influenced subsequent behaviour. Self-efficacy → performance research also draws on the definition of self-efficacy as a future-oriented, predictive construct: measures of self-efficacy involve statements of confidence in ability to achieve a future performance goal (Bong, 2012). This has practical implications, as research on this relationship draws impetus from (and feeds into) educational settings, in which interventions are sought to improve performance.

In a review of self-efficacy → performance research conducted over the past three decades, Klassen and Usher (2010) describe self-efficacy as a crucial and powerful influence on academic performance, accounting for approximately a quarter of the variance in outcomes, controlling for instructional variables. As such, self-efficacy is argued to rival previous performance and mental ability in its power to predict academic performance (Pajares & Kranzler, 1995). Meta-analyses of cross-sectional studies point to self-efficacy as one of the strongest and most consistent correlates of academic performance. For example, Richardson, Abraham, and Bond (2012) found that self-efficacy was by far the strongest correlate of tertiary GPA (ρ=.59), exceeding high school GPA (ρ=.41), tests of scholastic aptitude (ρ=.31–.33), and intelligence (ρ=.21). Similar moderate positive relationships are reported by Multon, Brown, and Lent (1991) and Honicke and Broadbent (2016).

Longitudinal research in the self-efficacy → academic performance paradigm is comparatively sparse (Honicke & Broadbent, 2016). Findings are consistent across school-aged and tertiary samples however, in that self-efficacy positively predicts subsequent academic performance, over periods ranging from a single semester, to courses over several years (rs=.37–.52; Chiang & Lin, 2014; Garriott & Flores, 2013; Majer, 2009; Parker, Marsh, Ciarrochi, Marshall, & Abduljabbar, 2014; Phan & Ngu, 2016).

Thus, it has long been suggested that direct manipulation of self-efficacy is a promising intervention strategy in learning settings (e.g., Bong, 2012; Pajares & Usher, 2008; Zimmerman, 1995). Given the practical implications for decision-making regarding educational reforms and interventions, as well as the implications for theory and research, it is important to be confident that the relationship between self-efficacy and academic performance is being interpreted accurately (Valentine, DuBois, & Cooper, 2004).

Academic performance → self-efficacy (I achieve; therefore I believe)

As the bulk of research on the relationship between self-efficacy and performance is cross-sectional, inferences about the direction of influence are impossible (Honicke &

Broadbent, 2016; Pajares & Usher, 2008). Some researchers argue that the cross-sectional relationships found are indicative of the influence of performance on subsequent self-efficacy, not the reverse (Shea & Howell, 2000; Vancouver et al., 2001), and some researchers even argue that self-efficacy is simply a proxy for past performance with no unique predictive power (Heggestad & Kanfer, 2005).

There is little doubt that performance outcomes influence self-efficacy. For example, previous mathematics performance positively predicts mathematics self-efficacy in both school-aged and tertiary samples (e.g., Klassen, 2004; Matsui, Matsui, & Ohnishi, 1990). Chin and Kameoka (2002) reported that reading scores predicted the educational self-efficacy of high school students (β=.32) after accounting for a range of demographic and psychosocial predictors. In a recent study, both standardised test scores and first semester GPA correlated with self-efficacy 12 months later, r=.30 (Lee, Flores, Navarro, & Kanagui-Munoz, 2015).

Such findings are not inconsistent with self-efficacy theory; in fact, investigations of the academic performance → self-efficacy relationship often grow out of self-efficacy theory directly. Mastery experience (an individual’s experience of performance success) is one of four posited sources of self-efficacy beliefs, along with vicarious experience, verbal social persuasion, and emotional physiological arousal (Bandura, 1997). Research shows that mastery experience is consistently the strongest (if not the only) predictor of the four possible sources (Britner & Pajares, 2006; Lent, Lopez, & Bieschke, 1991; Usher & Pajares, 2008).

While self-efficacy theorists do not discount the influence of performance on self-efficacy, they refute the argument that self-efficacy has no unique impact on performance (Bandura, 2012; cf. Heggestad & Kanfer, 2005). Some researchers have attempted to rule out the hypothesis that self-efficacy provides no incremental prediction of performance beyond that accounted for by previous performance, by undertaking meta-analyses of the self-efficacy → performance relationship, controlling for previous performance. Valentine et al.

(2004) used such an approach, interpreting the unique self-efficacy → performance effect (β=.10, *k*=9) as "small but noteworthy" (p. 127). Robbins et al. (2004) found that academic self-efficacy provided an incremental contribution to the prediction of academic achievement beyond that of socioeconomic status, standardised achievement measures, and high school GPA (β=.20, *k*=18).

The research summarised above provides evidence of both a self-efficacy → performance relationship and a performance → self-efficacy relationship; it also suggests that self-efficacy has an effect on subsequent performance which is incremental to that of previous performance alone. While this longitudinal research extends considerably on previous cross-sectional findings, it does not address the possibility that self-efficacy and performance are reciprocally related – nor does it elucidate the relative strength of directional effects.

A chicken-and-egg conundrum

The question of the direction of causality in the relationship between self-beliefs and academic performance has been described as one of "thorniest issues" in research in this area (Pajares & Schunk, 2001). While a great deal of research exploring academic self-efficacy and related constructs has been conducted since this assertion was made, questions regarding directional effects and reciprocity in the relationship between self-efficacy and performance are still moot. According to social cognitive theory, self-efficacy is embedded in a framework of reciprocal determinism, in which behaviour both shapes, and is shaped by a range of interacting factors (Bandura, 1977, 1997). In this model, self-efficacy and performance modify each other iteratively within a constant feedback loop (Multon et al., 1991): individuals are producers as well as products of their own cognitions, actions, and environments (Klassen & Usher, 2010). In educational settings, this means that learners reflect on their performance and use this information when formulating their self-efficacy beliefs, which then influence subsequent performance (Phan, 2012).

Several recent studies demonstrate increased interest in the mutual influences of self-efficacy and academic performance over time. In longitudinal studies of both high school and university students in which multiple measurements of self-efficacy and performance are staggered over several years, self-efficacy and performance predict each other, either in a self-efficacy → performance → self-efficacy pattern, or in a performance → self-efficacy → performance pattern (e.g., Caprara, Vecchione, Alessandri, Gerbino, & Barbaranelli, 2011; Hwang, Choi, Lee, Culver, & Hutchison, 2016). These types of designs provide evidence of positive mutual temporal effects, but by staggering data collection over time they do not enable the modelling of simultaneous reciprocal effects (Rogosa, 1988).

Thus, while reciprocal determinism between academic self-efficacy and academic performance may be considered a fait accompli (Lindsley, Brass, & Thomas, 1995), there is little direct empirical evidence which supports this proposition (Williams & Williams, 2010). This gap in the literature is likely due, in part, to the paucity of longitudinal studies in the area, a problem which is compounded because the two unidirectional research paradigms in this area are pursued largely independently of each other (Shea & Howell, 2000). In the most recent self-efficacy → performance meta-analysis, Honicke and Broadbent (2016) suggest a reciprocal relationship exists and recommend that this be investigated directly.

In the case of self-efficacy and performance, the issue of how to assess the unique influence of one variable on the other has been characterised by debate. Controlling for raw past performance (e.g., Feltz, 1982; Mitchell, Hopper, Daniels, George-Falvy, & James, 1994) is argued to be an over-correction (Bandura & Locke, 2003), while residualising self-efficacy from past performance (Bandura & Locke, 2003) is argued to lead to statistically artefactual results (Heggestad & Kanfer, 2005). In fact, as both past self-efficacy and past performance are expected to be covarying common-cause variables (Bandura, 2012; Feltz, Chow, & Hepler, 2008; Heggestad & Kanfer, 2005; Vancouver et al., 2001), any unidirectional approach will result in inflated estimates of the influence of these variables on each other (Brown et al., 2008) obscuring the properties of self-efficacy ↔ performance spirals (Lindsley et al., 1995; Shea & Howell, 2000). An approach is required where both self-efficacy and performance can be controlled at time 1 (Bandura, 2012; Feltz et al., 2008).

One approach that may overcome these limitations is cross-lagged panel analysis (CLPA; see Figure 1); a uniquely powerful approach to chicken-and-egg questions (Tyagi & Singh, 2014) which has been gaining traction in the behavioural sciences literature (e.g., Riketta, 2008). CLPA provides more robust evidence of potential reciprocal causality within dynamic relationships than other models, by meeting the requirement for temporal precedence, and demonstrating the unique effect of variable X on variable Y, and vice versa, controlling for other model paths (Selig & Little, 2012).

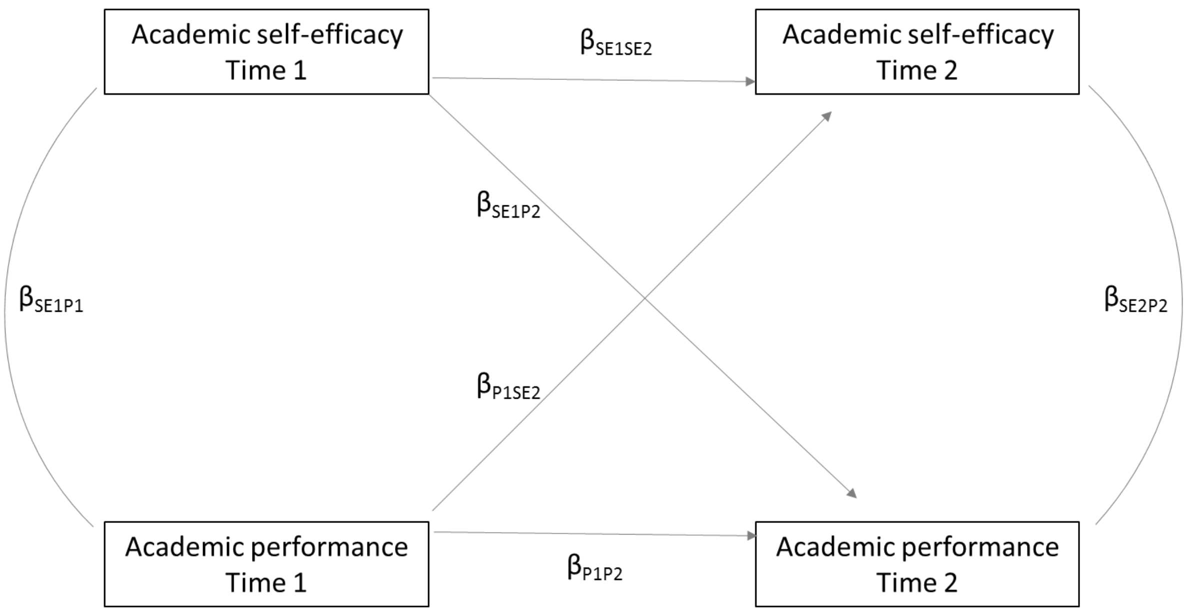


Figure 1. Conceptual diagram of the cross-lagged path model

Main aims of the present study

This study aims to explore the reciprocity and comparative strength of net directional effects in the self-efficacy ↔ academic performance relationship by meta-analysing existing longitudinal studies that provide panel data in a two-variable, two-wave design.

Six pooled correlations will be estimated and fit to a CLPA model (Figure 1) comprising two auto-correlations/stability coefficients (β_{SE1SE2} and β_{P1P2}), two synchronous or cross-sectional correlations (β_{SE1P1} and β_{SE2P2}), and two cross-lagged correlations (β_{SE1P2} and β_{P1SE2}). A reciprocal relationship is demonstrated by positive, significant cross-lagged path coefficients; the stronger antecedent, if one exists, would be marked by a larger coefficient.

A further aim of the present review is to assess variables which are likely to moderate the self-efficacy ↔ performance relationship. Moderator analyses are critical, given the concerns of researchers who consistently note that studies of the relationship between self-efficacy and academic performance are characterised by substantial heterogeneity (Honicke & Broadbent, 2016; Multon et al., 1991; Richardson et al., 2012; Valentine et al., 2004). In the following section, we identify participant- and study-level variables which may account for differences in effects across included studies. Further to accounting for heterogeneity in effects, these analyses contribute to our understanding of the conditions under which self-efficacy and academic performance are most strongly related – taking previous findings within a unidirectional self-efficacy → performance paradigm and extending these to the reciprocal self-efficacy ↔ performance relationship framework.

In summary, the main aims of this study are: to conduct a meta-analytic CLPA on studies presenting panel data regarding self-efficacy and academic performance, to test whether the data are consistent with reciprocal effects, to explore the relative antecedent strength of the two variables, and to test whether the relationship varies as a function of participant- and study-level variables identified as potential moderators in the literature.

Moderators of the self-efficacy ↔ performance relationship

Age. Participant age has been found to lead to differences in self-efficacy levels and/or variations in the relationship between self-efficacy and performance (e.g., Davis-Kean et al., 2008). The earliest meta-analysis of the relationship between self-efficacy and academic performance (Multon et al., 1991) reported a stronger correlation for college students (r=.35) than elementary students (r=.21). This difference likely relates to different levels of cognitive maturity, with adults more likely than children to use meta-cognitive strategies (Flavell, 1979) and reflect on self-constructs (Harter, 1999). Therefore, it is anticipated that the modelled relationships will be stronger for adults than for children.

Sex. Participant sex may also moderate the reciprocal relationship. Theoretically, differences may grow out of self-fulfilling prophecies whereby females are socialised to view themselves as less capable than males (Williams & Williams, 2010). Characteristics of educational settings may be implicated in this process, with associated differences in competence perceptions and performance (Eccles & Blumenfeld, 1985). Meta-analyses have reported sex differences in regard to academic performance (Richardson et al., 2012) and self-efficacy (Huang, 2013). However, Valentine et al. (2004) found that sex was not a moderator when self-efficacy was used to predict academic performance. Given the lack of consistency in findings, this moderator analysis will be exploratory in nature.

Self-efficacy scale type. The type of scale most appropriate for self-efficacy measurement is the subject of debate, and it is anticipated that scale differences will moderate findings. Bandura (2006) recommends the use of unipolar scales, with participants providing ratings of their degree of confidence from 0 – 100 percent in their ability to perform a future task, at each of several varying levels of difficulty of the task, moving from easy to hard. While the use of Likert scales is criticised (Klassen & Usher, 2010; Lee & Bobko, 1994) it remains very common (see Table 2). Unipolar scales are argued to be more sensitive and reliable (Bandura, 1986), while Likert scales are argued to be inappropriate for measuring self-efficacy, which does not have a positive/negative valence with a neutral midpoint (Bandura, 2012). Pajares, Hartley, and Valiante (2001) compared unipolar and Likert scales and found unipolar measures to be more psychometrically sound. Thus, it is hypothesised that models based on studies using unipolar scales will demonstrate stronger cross-lagged paths than those using Likert scales.

Lag between measurement waves. Variation in lag time between panel measurements is also considered as a moderating factor. In educational research, decisions about measurement timing often lack a strong rationale, resting instead on pragmatic considerations such as key academic dates (e.g., Phan, 2012), or conventions for ease of interpretation (e.g., Valentine et al., 2004). However, both theoretical and empirical considerations suggest otherwise: Shorter lags between measurement of self-efficacy and academic performance are likely to be associated with stronger effects (Bandura, 1997; Gist & Mitchell, 1992; Moriarty, 2014), potentially owing to the increasing likelihood that intervening variables will 'wash out' the relationship over time (Honicke & Broadbent, 2016; Valentine et al., 2004). Recent research findings appear to be consistent with this position (Galyon, Blondin, Yaw, Nalls, & Williams, 2012; Gore Jr, 2006; Obrentz, 2012; Zusho, Pintrich, & Coppola, 2003). Thus, we hypothesise that shorter lags will be associated with stronger effects than longer lags.

Specificity of measurement/match between specificity of self-efficacy and performance. Measurement issues regarding the degree of specificity of the self-efficacy measure and the degree of match between the self-efficacy and performance measures have been discussed for many years (Bandura, 1986). Further, these issues have been demonstrated to be critical to the accuracy of prediction of behaviour (Pajares & Usher, 2008). Domain-based self-efficacy measures reflect confidence in one's ability to perform academically in a subject or particular course; e.g., I'm sure I can do even the hardest work in my math class

(Lewis et al., 2012). Domain-based self-efficacy can also be measured by asking individuals to rate their levels of confidence in achieving a range of grades on a subject-specific course (e.g., Beck & Schmidt, 2013). Task-based self-efficacy measures reflect confidence in one's ability to perform particular academic tasks; often, this will be answering specific questions correctly (e.g., Bonne & Johnston, 2016) or attaining a particular grade on an isolated performance task such as an exam (e.g., Mone, Baker, & Jeffries, 1995). It is generally reported that more specific self-efficacy measures are more strongly correlated with performance than less specific measures, and more closely matched measures of self-efficacy and performance are more strongly correlated with each other (Bong, 2001; Pajares & Miller, 1995; Zimmerman, 1995). In addition, reviewers report that measurement specificity (Multon et al., 1991) and match (Valentine et al., 2004) moderate the unidirectional relationship between self-efficacy and performance in meta-analyses. Thus, it is hypothesised that models based on studies using more specific self-efficacy measures, and matched specificity of self-efficacy and performance, will demonstrate stronger model paths than those with less specific measures and non-matched specificity.

Cohort effects. We note that, as no restrictions were put in place in terms of date of article publication, it is possible that cohort effects may be found in the sample. Broad changes in educational policies and assessment practices over time may affect the self-efficacy ↔ performance relationship (Goldstein, 1983). An exploratory moderator analysis using time since publications as a variable will be undertaken to test this possibility.

Summary

By using meta-analytic cross-lagged panel analysis, this systematic review addresses the chicken-and-egg question of self-efficacy and performance, bringing together two avenues of enquiry which have so far been largely isolated (self-efficacy → performance *versus* performance → self-efficacy). To our knowledge, a meta-analytic cross-lagged panel analysis is yet to be conducted with the constructs of self-efficacy and academic performance.[2] The present review builds on others (e.g., Richardson et al., 2012; Valentine et al., 2004) in several important ways. Firstly, the present analysis assesses the net effect of self-efficacy on performance while holding all CLPA model paths constant, rather than just the effect of previous performance. Secondly, this is the first analysis, to our knowledge, of the net effect of academic performance on subsequent academic self-efficacy, accounting for other model paths including initial levels of self-efficacy. Finally, our approach applies stricter inclusion and exclusion criteria than previous reviews in order to ensure adequate construct validity and obtain a sample of more homogeneous studies that will allow testing the proposed relationships with more reliability than previous studies.

A better understanding of the complexities of the relationship between self-efficacy and academic performance will provide valuable input into theory development and research design, and will also help to inform education policy and teacher training programs.

Method

This systematic review and meta-analysis was undertaken in accordance with the PRISMA guidelines (Moher, Liberati, Tetzlaff, & Altman, 2009).

Inclusion criteria

Operationalisation of self-efficacy. There is evidence that measurement issues pose a genuine threat to the validity of research in this area (Bandura, 2012; Klassen & Usher, 2010). Previous reviews have accounted for heterogeneity across studies by differences in the operationalisation of both self-efficacy and performance (Honicke & Broadbent, 2016; Multon et al., 1991; Richardson et al., 2012; Valentine et al., 2004). Self-efficacy is often not

[2] It should be noted that, in a meta-analysis of cross-lagged panel studies of the broader self-efficacy/performance relationship by Sitzmann and Yeo (2013), a subset of studies were related to self-efficacy and academic performance. However, the methodological rigour of this study and its overall contribution to the literature has been questioned (Bandura, 2015). For example, it has been noted that some included studies did not measure one of the two key constructs (Bandura, 2015).

appropriately measured, nor differentiated from other constructs such as self-esteem and self-concept (Bandura, 2012; Klassen & Usher, 2010; Lee & Bobko, 1994; Pajares, 1996). This problem is compounded when reviews cast a wide net, because very broad definitions of variables may result in the conglomeration of distinct constructs (Bandura, 2012). To address these issues, it has been recommended that researchers use more narrowly defined and theoretically grounded measures – which we will apply in this present review (Klassen & Usher, 2010; Robbins et al., 2004; Valentine et al., 2004). We include studies when scales reflect the three key components of self-efficacy: a personal judgement of ability (internal attribution) to prospectively (future orientation) perform academically (a behaviour, as opposed to an attitude or personal characteristic) (Bandura, 2006; Schwarzer & McAuley, 2016). Global/generalised/"trait" measures will not be included, owing to theoretical concerns about their applicability and their limited power to predict academic performance (Bandura, 2012; Klassen & Usher, 2010; Pajares, 1997). Furthermore, this review does not include efficacy for self-regulated learning, e.g., *How well can you arrange a place to study without distractions?* (Zimmerman, Bandura, & Martinez-Pons, 1992). In sum, we seek measures which reflect a can do judgement for future academic performance (Table 1).

Operationalisation of academic performance. Previous reviews have also suffered from inconsistent operationalisations of academic performance (Robbins et al., 2004). Multon et al. (1991) and Honicke and Broadbent (2016) used a wide range of performance metrics in their reviews, and both reported that performance specification moderated the relationships found. A more precise definition of academic performance would contribute to the validity of analyses of the self-efficacy/performance relationship. While GPA is a readily accessible measure of academic performance, cumulative measures are not considered appropriate because aggregation potentially blurs performance across subject areas, task types, and – most importantly for the present analysis – across time. To provide precision of measurement and avoid the pitfalls of aggregated measures, we define academic performance as objective scores on individual performance tasks, such as tests or exams. Such a definition is also consistent with the recommendations of self-efficacy theorists.

With the above rationale in mind, studies were included if they:

1. Presented data regarding self-efficacy (academic/domain/task), measured in accordance with self-efficacy theory (see Table 1).
2. Collected data regarding academic performance on a specific occasion.
3. Collected data on self-efficacy and performance using a panel design in accordance with the path model shown in Figure 1, providing a set of six correlations.
4. Included participants in traditional primary, secondary and tertiary educational settings.
5. No interventions were conducted between the measurement waves, as interventions may obscure the effects of self-efficacy and academic performance on each other. Control group data were eligible for inclusion if other inclusion criteria were met.

We note that, in a single measurement wave, some researchers measure self-efficacy immediately prior to performance (e.g., Galyon et al., 2012), while some measure self-efficacy in the days or weeks before the performance measurement (e.g., Richard, Diefendorff, & Martin, 2006). Often, this is for practical reasons: to increase participation (Beck & Schmidt, 2013), or to reduce effects of exposure to actual test questions in a self-efficacy measure (Bonne & Johnston, 2016). Research indicates that there is little difference in the prediction of subsequent performance whether self-efficacy is measured immediately prior or up to two weeks prior to performance (Beck & Schmidt, 2013; Mone et al., 1995). Thus, studies in which self-efficacy is measured within two weeks of the performance measurement were included, with different degrees of concurrency analysed as a potential risk of bias (section 2.6).

Table 1

Examples of self-efficacy measures: specificity and inclusion/exclusion

Excluded	Included - task	Included - domain	Included - academic
"I am confident of my own decision." (Hwang et al., 2016)	Confidence ratings for ability to achieve a range of grades on an impending exam. (Mone et al., 1995)	"I'm sure I can do even the hardest work in my math class." (Lewis et al., 2012)	"I'm confident I can master the courses I'm taking this semester." (Bong, 2001)
"I am a fast learner" (Richardson et al., 2012)	Confidence ratings for ability to correctly answer specific questions on an impending exam (Bonne & Johnston, 2016)	"I can explain the facts, concepts and arguments covered in this course to others in my own words" (Galyon et al., 2012)	"I'm certain I can figure out how to do the most difficult class work." (Hornstra, Van Der Veen, Peetsma, & Volman, 2013)
"I am good at maths." (Scott, 2000)	Confidence ratings for ability to perform particular writing tasks (Meier, McCarthy, & Schmeck, 1984)	Confidence ratings for achieving a range of course grades (Beck & Schmidt, 2013)	"How well can you pass all of your subjects in school?" (Galla et al., 2014)

Literature search

Electronic searches were conducted in Scopus, PsycINFO, the Education Resources Information Centre (ERIC) and Web of Science. Three search fields covered terms relating to self-efficacy, academic performance (e.g., school, education, test), and methodological terms (e.g., panel, longitudinal, lag). Full search parameters are appended (Appendix 2.1). In order to reduce the risk of publication bias, unpublished papers such as dissertations and theses were eligible for inclusion (Smith, 1980). In addition, there were no restrictions on year of publication or language, as long as abstracts were available in English.

A flow chart of the search is appended (Appendix 2.2). The final search was conducted in April 2016. In total, 5,487 studies were located in the database search. With duplicates removed, 3,177 articles remained for analysis. After screening titles and abstracts for relevance, 197 articles remained for potential inclusion. We manually searched the reference lists of known extant meta-analyses which included analyses of the relationship between self-efficacy and academic performance. Where articles provided summaries of included studies (Honicke & Broadbent, 2016; Robbins et al., 2004; Sitzmann & Yeo, 2013; Valentine et al., 2004), these were used to identify studies which measured both self-efficacy and performance and had a longitudinal design, with 36 studies identified for potential inclusion. In other cases (Multon et al., 1991; Richardson et al., 2012), all titles of included studies were screened for relevance (n=252) and a further 114 potentially relevant studies were identified. Full texts of these 347 articles were then sought out and further screened according to the inclusion criteria. If there was any uncertainty in regard to whether studies met inclusion criteria (e.g., if self-efficacy scale items were not included in the full text and were otherwise unpublished, or if the timing of data collection at any given measurement wave was not explicitly stated) authors were contacted for this information.

Twenty-three records, about half of which were dissertations, which had been identified for full-text analysis were excluded on the basis that the required data were unavailable after multiple attempts to contact the relevant authors. The vast majority of exclusions were made either because studies did not measure self-efficacy or performance (either at all, or in accordance with the specifications above) or because studies did not collect data in a panel design. Close examination of study methodologies was necessary as some studies superficially appeared to be panel designs (from their use of terms such as "time 1" and "time 2" or "wave 1" and "wave 2"), but on consultation with authors, the variables were actually measured some months apart at a given measurement "wave". Three studies were excluded on the basis of measuring self-efficacy subsequent to performance, rather than prior to performance, at each measurement wave (Moriarty, 2014). Ultimately, 11 studies were identified which met the inclusion criteria.

Data extraction

The first author used coding sheets to extract the required data from included studies. Correlation matrices were extracted from articles where available (k=4), or obtained from authors (k=7). In order to ensure independence of data points, and to avoid difficulties interpreting three or more waves of data in panel analysis (Rogosa, 1988), each study provided a single set of six correlations as per Figure 1. Several studies reported additional measurement waves; in order to fit data to the CLPA model, only the first two waves were included. Several studies measured self-efficacy at multiple levels of specificity, in which case, the measure which most closely matched the performance measure was used, as such a match is believed to result in the greatest predictive power (Klassen & Usher, 2010; Pajares, 1997). Coding also covered sample size, publication status, study location, participant variables and study characteristics relating to moderator analyses (section 2.5) and factors potentially associated with risk of bias (section 2.6). A subset of included studies was coded by KN and inter-rater agreement was 100%, owing to the transparency of coding categories.

Overall analysis

Using Metafor for R (Viechtbauer, 2010), six pooled correlations, one for each path in the cross-lagged panel model, were extracted from the zero-order correlations provided by the included studies. The raw correlations were subjected to a Fisher's z transformation, and random effects modelling was used as the included studies are believed to represent a distribution of true effects (Borenstein, Hedges, Higgins, & Rothstein, 2009a). Restricted maximum likelihood estimation was used to ensure accurate estimation of variance (Thompson & Sharp, 1999), and study weighting was based on the inverse of study-specific variance, giving more weight to studies with greater precision (Viechtbauer, 2010). Analyses were conducted on the zero-order correlations using *p*-curve 4.05 (Simonsohn, Nelson, & Simmons, 2015) to calculate the underlying statistical power for the estimation of each of the six pooled correlations. The matrix of pooled correlations resulting from the above analysis served as input for the CLPA (Figure 1) using MPLUS 7 (Muthen & Muthen, 1998-2012).

Moderator analyses

The overall approach to moderator analyses was to test whether any of the six pooled correlations forming the basis of the CLPA varied as a function of the identified moderator variables. For moderators identified as significant, separate path models for subgroups of the moderator variable were estimated using the same procedure as for the overall analysis (section 2.4). Subsequent to the estimation of the separate path models for moderator subgroups, the cross-lagged path coefficients for contrasting subgroups were tested for significant differences using the pooled standard deviation method (Kock, 2014).

Tests of the effects of proposed moderators were undertaken using meta-regression analyses in R using the metafor package (Viechtbauer, 2010). Within the framework of random-effects meta-analysis, meta-regression is an analysis of the relationship between study-level variables and effect size – analogous to multiple regression in primary studies assessing the relationship between subject-level variables and an outcome (Borenstein, Hedges, Higgins, & Rothstein, 2009b). As a significant moderator effect for any individual pooled correlation could result in changes to the overall meta-analytic path analysis, moderator analyses were conducted for any moderator that showed significant Q statistics for any of the six individual model paths. R^2 values, which indicate the amount of variance in the pooled correlations accounted for by the moderator variable, were also calculated.

Moderator specification. Continuous values for moderators were available for sex (proportion of men; data was unavailable for 2 studies), cohort (years since publication was used as a proxy as the specific timing of data collection was not published in most cases), and time lag between waves (in weeks). The nature of the data necessitated the use of categorical moderator variables in several cases as follows.

Age. Average participant age was not specified in more than half of the studies included in this review, so a categorical moderator was constructed whereby participants were defined as either children or adults, using age data, where available, or participant education level information (e.g., grade 5 versus tertiary education]). In this way, all studies could be included in the moderator analysis.

Scale. A categorical moderator was coded such that unipolar scales were compared with Likert scales. In one case, the scale could not be identified as either unipolar or Likert (Vancouver & Kendall, 2006). As this single-item scale was considered to be psychometrically less sound than a unipolar scale (Bandura, 2012), this study was grouped with the Likert scales.

Specificity/match. Self-efficacy was coded as either domain-based or task-based. In the included studies, all measures of performance were incidentally task-based (e.g., exams, tests). This rendered coding for both self-efficacy specificity and match between self-efficacy and performance measures unnecessary, because these groups overlapped completely. That is, because performance was task-based, studies with task-based self-efficacy were characterised both as being matched to performance and as having a more specific level of self-efficacy, whereas studies with domain-based self-efficacy measures were characterised both as being unmatched to performance and as having a less specific level of self-efficacy. Therefore, only one moderator analysis was required, and for ease of interpretation, we refer hereafter to the moderator categories as “matched” or “non-matched”.

Assessment of risk of bias

Risk of bias was assessed taking into account the reliability of self-efficacy and performance measurements, the concurrency issue mentioned in section 2.1, publication/selective reporting bias, and author bias. We note that some potential sources of bias (e.g., atheoretical or imprecise operationalisations of constructs) are minimised through the application of the specific inclusion criteria outlined in section 2.1.

Reliability of self-efficacy measurement. All reported Cronbach’s alpha reliability coefficients for self-efficacy measures were acceptable at >.70, with all but one exceeding α=.83 (Nunnally & Bernstein, 1994). Nonetheless, self-efficacy reliability was analysed as a potential continuous moderator of the six pooled correlations, using the same meta-regression approach as was taken with the moderators (section 2.5). None of the model paths were significantly moderated by reliability (all Q statistics, $p>.05$); therefore, reliability of the self-efficacy measure is not deemed to present a risk of bias in the present study.

Reliability of performance measurement. Only three studies provided reliability information for the performance measure; therefore, a moderator analysis was not conducted. All performance measures were taken from academic records and are therefore deemed to be at low risk of bias.

Concurrency of measurement. As mentioned above, studies were defined as meeting the requirements for panel measurement when self-efficacy and performance were measured within two weeks of each other at any given measurement wave. In order to assess the validity of this approach to inclusion, a categorical moderator variable was coded such that those studies in which all measures were taken in the same sitting were compared with all other studies. This variable did not significantly moderate any of the pooled correlations that served as input for the path model (all *Q* statistics, $p>.05$). Therefore, differences in concurrency of measurement are not deemed to present a risk of bias in the current study.

Selective publication/reporting bias. We screened both published and unpublished papers in the literature search process, and conducted both traditional (Failsafe N) and contemporary (p-curve) analyses of the risk of selective publication/reporting bias. Results (see section 3) suggest minimal risk of bias based on selective publication/reporting.

Author bias. Inclusion criteria and coding procedures were transparent and unambiguous, resulting in 100% inter-rater agreement on coding in a subset of studies; thus, the current meta-analysis is deemed to be at low risk of author bias.

Results

Included studies (k=11) were published between 1984 and 2016. Study sample sizes ranged from 50 to 1,456 (M=244). Ten studies were conducted in North America, with one study conducted in New Zealand. There were ten published articles and one conference paper. Ten studies comprised mostly white participants; a single study comprised mostly Hispanic participants. Six studies considered mathematical domains, while the remainder considered management, writing, human development, and psychology. Further details on features of included studies, including zero-order correlations and moderator information, are in Table 2.

Overall analysis

Table 3 shows the pooled estimates of the six correlations in the cross-lagged panel model, along with standard errors, and 95% confidence intervals. Also in Table 3 are the results of tests of heterogeneity, selective publication/reporting bias, and observed power. The failsafe *N* values indicate that substantial numbers of additional null-effect studies would be needed to increase the *p*-value for each path to more than .05 (Rosenthal, 1979). Data sets which are subject to selective reporting or *p*-hacking are likely to show a left-skewed *p* distribution. The full- and half- *p*-curve analyses all show right-skewed distributions, indicating that the data included in the calculation of each of the six model paths have evidential value (Simonsohn, Simmons, & Nelson, 2015)[3]. Observed power estimates, ranging from 91% to 99%, are also reported in Table 3. The I^2 and *Q* heterogeneity test values suggest sampling error alone is unlikely to account for the differences in effect sizes. Forest and funnel plots for each of the six paths are appended (Appendices 2.3 and 2.4). Overall, the pooled correlations are consistent with previous findings in terms of direction and magnitude, with moderate-to-strong positive auto-correlations, and small-to-moderate positive synchronous and cross-lagged correlations.

[3] The *p*-curve user guide directs authors to present a *p*-curve disclosure table. Such a table is most applicable to syntheses of varied experimental data. In the present case, much of the information that would be presented in such a table is identical across studies (e.g., all designs and statistics are correlational). Other relevant information which would form part of the disclosure table (e.g., statistical values) is in Table 2.

Table 2

Features of included studies including zero-order correlations and moderators

Study	n	Age	% m	SE α	CC	Lag	Match	Scale	SE1P1	SE1P2	SE2P1	SE2P2	SE1SE2	P1P2
Beck et al., 2013	69	A	ns	0.94	No	S	No	U	0.333	0.292	0.488	0.313	0.736	0.697
Bonne et al., 2016	50	C	36	ns	No	L	Yes	L	0.570	ns	0.370	0.310	0.300	0.390
Daniels et al., 2010	111	A	ns	0.83	Yes	S	No	U	0.201	0.229	0.429	0.219	0.676	0.541
Finney et al., 2003	103	21	32	0.91	Yes	L	Yes	L	0.480	0.224	0.404	0.441	0.314	0.340
Galyon et al., 2012	213	A	26	0.92	Yes	S	Yes	L	0.079	0.026	0.239	0.213	0.547	0.570
Lewis et al., 2012	1456	C	47	0.84	No	L	No	L	0.318	0.272	0.285	0.353	0.399	0.680
Meier et al., 1984	71	A	57	0.85	Yes	L	Yes	U	0.361	0.370	0.210	0.090	0.690	0.460
Mone, 1994	252	21.6	59	0.87	Yes	S	Yes	U	0.310	0.290	0.480	0.320	0.670	0.490
Mone et al., 1995	215	A	53	0.85	No	S	Yes	U	0.320	0.220	0.450	0.240	0.650	0.530
Richard et al., 2006	83	19.9	21	0.83	No	S	No	L	0.340	0.340	0.480	0.390	0.760	0.630
Vancouver et al., 2006	65	21	21	0.83	No	S	Yes	O	0.108	0.253	0.405	0.319	0.630	0.189

Note: Mean age is provided where reported, otherwise, age group (adults=A, children=C) is provided based on sample information. SE α = Cronbach's alpha for self-efficacy; %m = proportion males; CC = concurrency; L=long, S=Short; U=unipolar, L=Likert, O=Other; ns = not specified; SE1P1 = self-efficacy time 1 with performance time 1, etc.

Table 3

Pooled correlations and tests of heterogeneity, publication bias and observed power: Overall analysis

	k	n	SE1P1	SE1P2	SE2P1	SE2P2	SE1SE2	P1P2
Pooled correlation	11	2688	0.316***	0.248***	0.402***	0.312***	0.689***	0.583***
(SE)			(0.04)	(0.03)	(0.04)	(0.03)	(0.07)	(.06)
95%CI			.231, .401	.182, .313	.328, .477	.255, .369	.551, .827	.468, .697
I^2			68.15	45.32	58.61	30.95	88.89	83.42
Q			26.007**	14.777	26.5777**	14.1855	191.9659***	71.3827***
Failsafe N			749	431	1144	735	3144	2979
p-curve, full (Z)			-9.56***	-6.02***	-13.31***	-7.98***	-19.99***	-17.58***
p-curve, half (Z)			-9.63***	-5.26***	-12.65***	-7.28***	-20.68***	-17.08***
Observed power			99%	94%	99%	91%	99%	99%
[90%CI]			[98%, 99%]	[77%, 99%]	[99%, 99%]	[91%, 99%]	[99%, 99%]	[99%, 99%]

Note: Correlations are Fisher's Z transformations. SE1P1 = self-efficacy at time 1 with performance at time 1; SE1P2 = self-efficacy at time 1 with performance at time 2; SE2P1 = self-efficacy at time 2 with performance at time 1; SE2P2 = self-efficacy at time 2 with performance at time 2; SE1SE2 = self-efficacy at time 1 with self-efficacy at time 2; P1P2 = performance at time 1 with performance at time 2. **p<.01 ***p<.001

Figure 2 displays the meta-analytic path model for the overall analysis, estimated from the pooled correlations in Table 3. All of the paths are positive and statistically significant, $p<.001$. Focusing on the cross-lagged paths, the performance → self-efficacy effect was β=.205, whereas the self-efficacy → performance effect was β=.071. These estimates are based on a fully saturated model. A model was also tested in which the cross-lagged paths were constrained to be equal, indicating an equivalent reciprocal relationship. A Chi-squared test of the constrained model was interpreted as a Chi-squared difference test, comparing whether the model constraining the cross-lagged paths to be equal resulted in a loss of fit compared to the model allowing them to vary: if this test was significant, the constrained model provided a significantly poorer fit than the saturated model (Menard, 2007). In this case, the constrained model provided a significantly poorer fit to the data, $\chi^2 = 37.47$, $p<.0001$; thus, the cross-lagged paths are significantly different from each other.

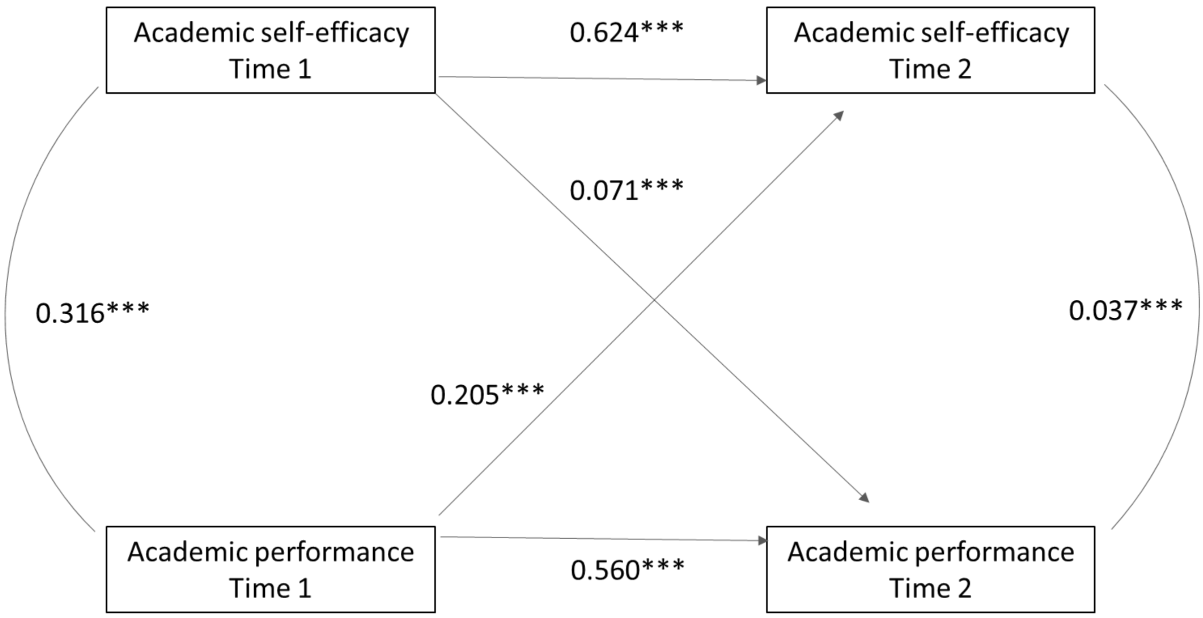


Figure 2. Cross-lagged panel model: overall analysis

Note: Figures are standardised path coefficients. A model in which the cross-lagged paths were constrained to be equal provided a significantly poorer fit to the data (see section 3.1). ***$p<.001$

Overall, these results are consistent with a reciprocal effects model: self-efficacy exerts a unique influence on subsequent performance, and performance has a net influence on subsequent self-efficacy. Performance is identified as a significantly stronger antecedent of self-efficacy than the reverse, with the performance → self-efficacy effect size almost three times the size of that of self-efficacy → performance.

Moderator analyses

The relationship between academic self-efficacy and academic performance was significantly moderated by sample age, length of the time lag, the degree of match between the self-efficacy measure and the performance measure, and the type of self-efficacy scale used. The proportion of male participants and elapsed time since publication did not moderate the relationship. Complete meta-regression analyses are appended (Appendix 2.5).

Path analyses were conducted on eight separate pooled correlation matrices, with two subgroups for each of the four significant moderator variables. For those variables which were already categorical, the subgroups as specified in in section 2.5.1 were used for the separate path analyses (i.e., adults and children, matched and non-matched specificity, and Likert and unipolar scales). Time lag had been analysed as a continuous moderator; thus, we formed short and long categories for comparison.[4]

Pooled correlations, standard errors and 95% confidence intervals for the moderator categories are shown in Table 4, with a similar pattern of correlations to the overall model. Heterogeneity test values, observed power, forest plots and funnel plots for pooled correlations for moderator subgroups are appended (Appendices 2.6 and 2.7).

---

[4] For this categorisation, we note that recent findings in the literature indicate that self-efficacy is more strongly related to performance when measured up to 2 months prior to the performance measure, with weaker relationships associated with longer lags (from 3 months upwards) (Gore Jr, 2006; Obrentz, 2012; Phan, 2012; Zusho et al., 2003). In the included data set, this lag range coincided with a natural division in lag times: 7 studies had lag times of up to 7 weeks, and 4 studies had longer lags of 3 to 12 months, while no studies had a lag of between 7 weeks and 3 months. This categorisation therefore provided theoretically grounded as well as readily interpretable differentiation of studies with a reasonable number of studies in each category.

Table 4

Pooled correlations: Moderator analyses

		n		SE1P1	SE1P2	SE2P1	SE2P2	SE1SE2	P1P2
Age group	Adults	1182	r	.290***	.240***	.427***	.288***	.758***	.562***
			(SE)	(0.05)	(0.04)	(0.04)	(0.03)	(0.06)	(0.06)
			95%CI	.196, .383	.161, .320	.345, .509	.226, .351	.631, .885	.448, .676
	Children	1506	r	.456**	.279***	.296***	.367***	.419***	.645**
			(SE)	(0.16)	(0.02)	(0.03)	(0.03)	(0.03)	(0.21)
			95%CI	.151, .761	.228, .330	.245, .347	.317, .418	.368, .470	.239, 1.051
Lag	Short	1008	r	.249***	.229***	.448***	.282***	.797***	.598***
			(SE)	(0.05)	(0.05)	(0.05)	(0.03)	(0.05)	(0.07)
			95%CI	.158, .341	.137, .320	.357, .539	.220, .345	.704, .889	.470, .726
	Long	1680	r	.433***	.281***	.301***	.333***	.473***	.544***
			(SE)	(0.07)	(0.02)	(0.02)	(0.07)	(0.12)	(0.12)
			95%CI	.295, .571	.232, .329	.253, .349	.195, .470	.241, .705	.309, .778
Matched specificity	Yes	969	r	.328***	.223***	.397***	.285***	.643***	.487***

			(SE)	(0.07)	(0.05)	(0.05)	(0.04)	(0.08)	(0.06)
			95%CI	.187, .469	.120, .327	.298, .495	.210, .359	.484, .802	.377, .597
	No	1719	r	.323***	.281***	.420***	.360***	.779***	.773***
			(SE)	(0.02)	(0.02)	(0.07)	(0.02)	(0.14)	(0.06)
			95%CI	.276, .371	.233, .328	.285, .554	.312, .407	.509, 1.05	.664, .882
Scale	Uni-	718	r	.314***	.275***	.473***	.265***	.818***	.592***
	polar		(SE)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
			95%CI	.240, .388	.201, .349	.399, .547	.191, .339	.744, .892	.514, .671
	Likert	1970	r	.328***	.221***	.346***	.349***	.566***	.546***
			(SE)	(0.09)	(0.06)	(0.04)	(0.04)	(0.11)	(0.10)
			95%CI	.160, .496	.104, .338	.261, .431	.276, .422	.358, .774	.346, .747

Note: r = pooled correlations after Fisher's Z transformations; SE= self-efficacy; SE1P1 = self-efficacy at time 1 with performance at time 1; SE1P2 = self-efficacy at time 1 with performance at time 2; SE2P1 = self-efficacy at time 2 with performance at time 1; SE2P2 = self-efficacy at time 2 with performance at time 2; SE1SE2 = self-efficacy at time 1 with self-efficacy at time 2; P1P2 = performance at time 1 with performance at time 2; *p<.05 **p<.01 ***p<.001

Figures 3 to 6 show the path models for the subgroups of moderator variables found to significantly impact on the relationship between self-efficacy and academic performance, estimated on the basis of the pooled correlations in Table 4. These models were largely consistent in that most paths were again positive and statistically significant. One key exception is that, for children, the β_{SE1P2} cross-lagged path was near-zero and did not reach statistical significance. In all cases, the performance → self-efficacy paths were stronger than the corresponding self-efficacy → performance paths.

As with the overall analyses, Chi-squared tests were conducted to ascertain whether improved model-fit compared to saturated models was obtained with cross-lagged paths constrained to be equal. With the exception of long time lags (p=.07), models in which the cross-lagged paths were constrained to be equal provided a significantly poorer fit to the data (p<.001). Full details of these tests are appended (Appendix 2.8).

Table 5 shows the results of tests of differences between cross-lagged path coefficients for moderator subgroups. The reciprocal relationship was significantly stronger for adults than for children, for both of the cross-lagged paths. Studies with shorter lags showed stronger reciprocal effects than those with longer delays between measurement waves; this difference was statistically significant for the β_{P1SE2} path. Reciprocal effects were stronger for studies in which the specificity of the self-efficacy measure matched the performance measure, though the differences did not reach statistical significance. Studies with unipolar self-efficacy scales showed stronger reciprocal effects than those with Likert scales; this difference was statistically significant for the β_{P1SE2} path.

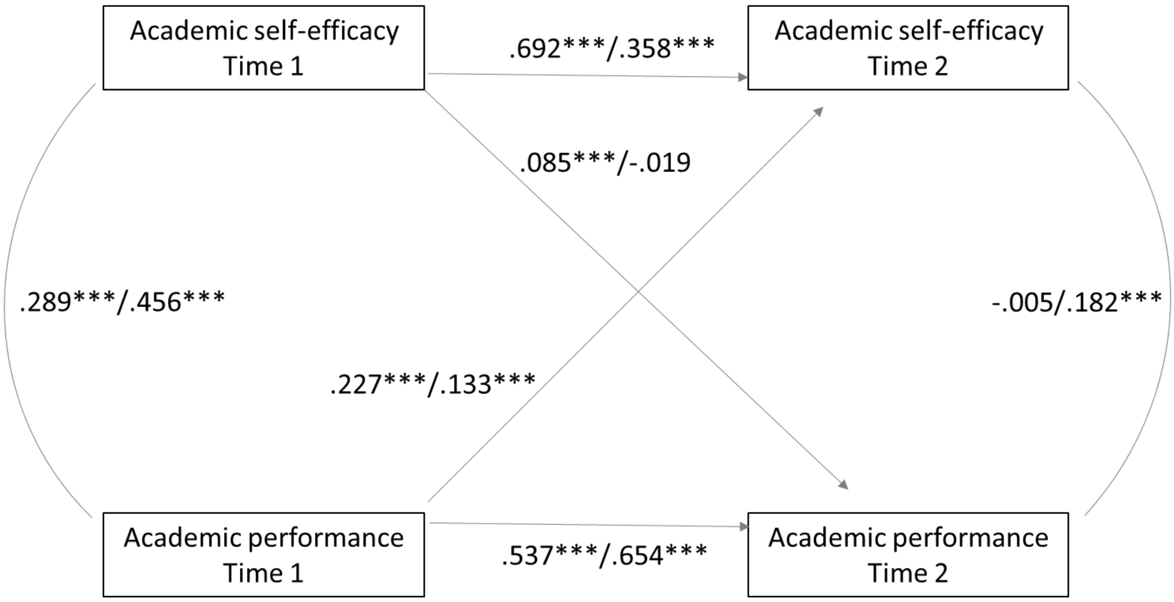


Figure 3. Cross-lagged panel models for adults/children respectively

Note: Figures are standardised path coefficients. For both adults and children, a model in which the cross-lagged paths were constrained to be equal did not provide a significantly poorer fit to the data (see Appendix 2.8). ***p<.001

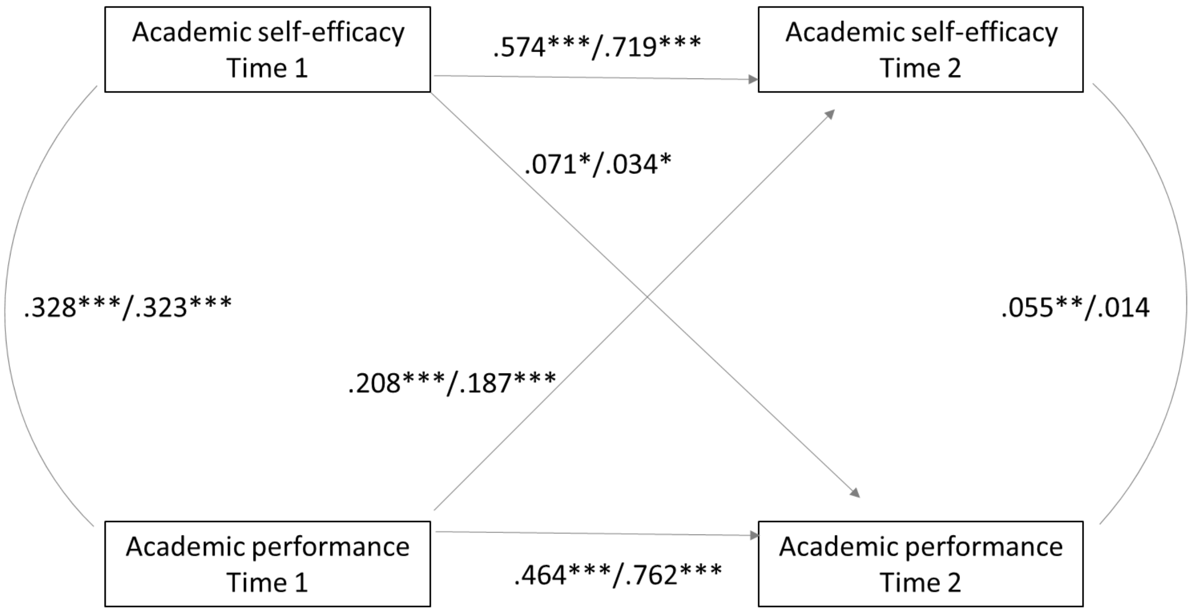


Figure 4. Cross-lagged panel models for matched/non-matched specificity respectively

Note: Figures are standardised path coefficients. For both matched and non-matched specificity, a model in which the cross-lagged paths were constrained to be equal provided a significantly poorer fit to the data (see Appendix 2.8). ***p<.001 **p<.01 *p<.05

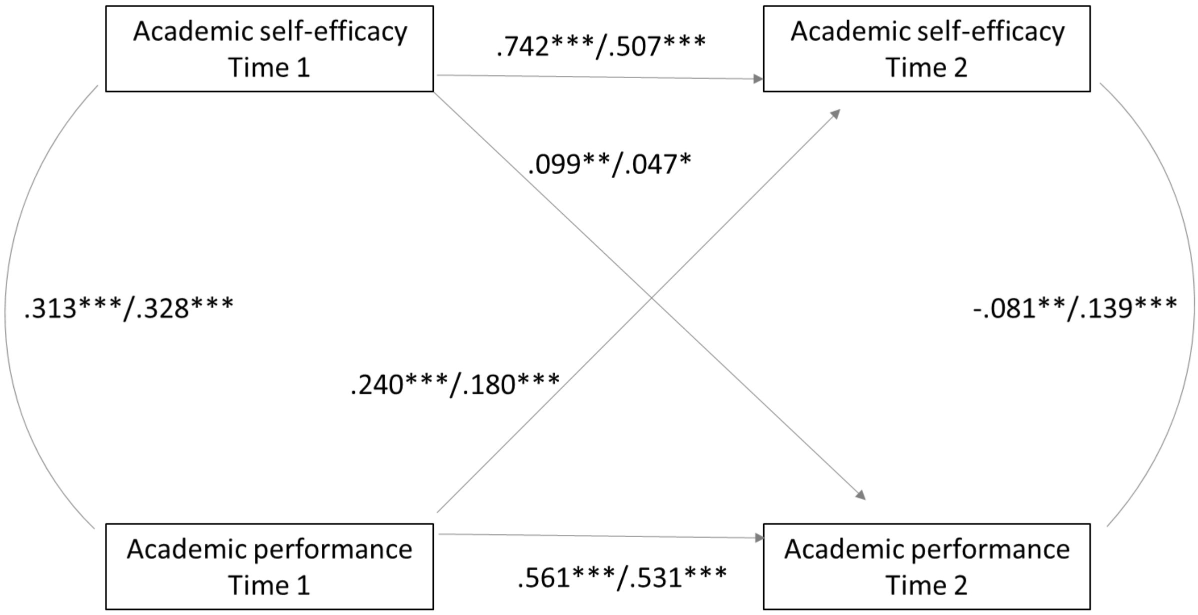


Figure 5. Cross-lagged panel models for unipolar/Likert self-efficacy scale respectively

Note: Figures are standardised path coefficients. For both unipolar and Likert scales, a model in which the cross-lagged paths were constrained to be equal provided a significantly poorer fit to the data (see Appendix 2.8). ****p*<.001 ***p*<.01 **p*<.05

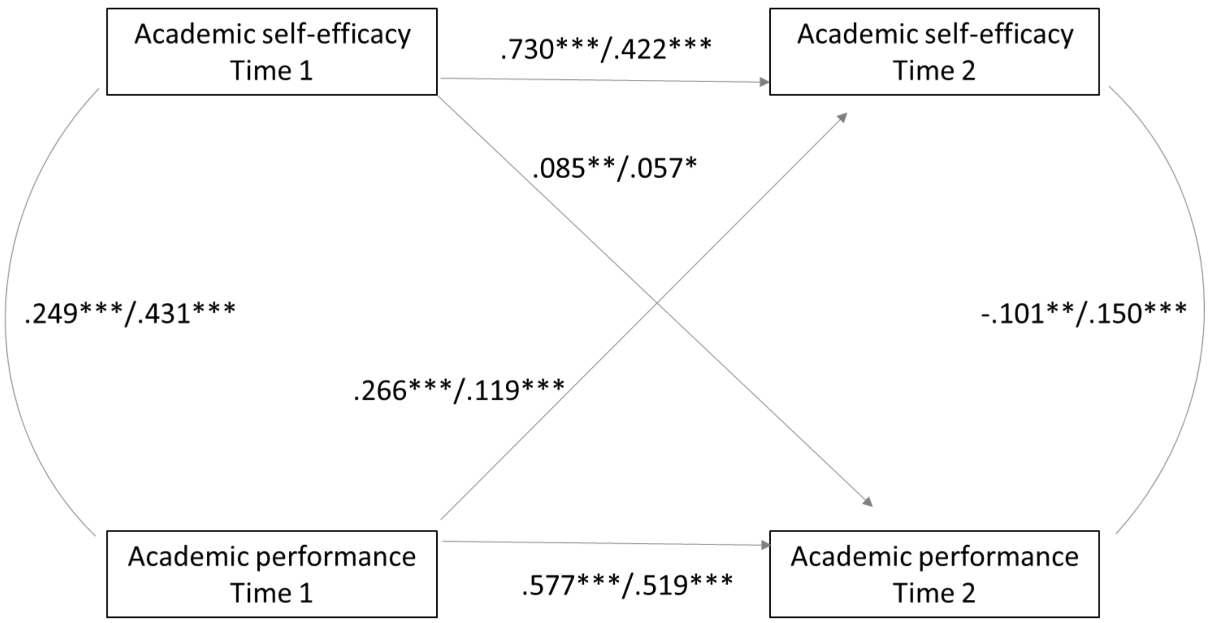


Figure 6. Cross-lagged panel models for short/long lag respectively

Note: Figures are standardised path coefficients. For the short lag, a model in which the cross-lagged paths were constrained to be equal provided a significantly poorer fit to the data (see Appendix 2.8). For the long lag, a model in which the cross-lagged paths were constrained to be equal did not provide a significantly poorer fit to the data (see Appendix 2.8). ****p*<.001 ***p*<.01 **p*<.05

Table 5

Tests of differences between cross-lagged path coefficients for moderator subgroups

	path	t	p
Age group	SE1P2	3.13	.0018
	P1SE2	3.71	.0004
Lag	SE1P2	0.78	.4200
	P1SE2	4.33	<.0001
Match	SE1P2	1.19	.2327
	P1SE2	.77	.4435
Scale	SE1P2	1.37	.1717
	P1SE2	1.77	.0342

Discussion

Overall analysis: I believe therefore I achieve, and vice versa

This study explored the chicken-and-egg conundrum in the relationship between academic self-efficacy and academic performance by means of a systematic review and meta-analytic CLPA. The meta-analysis provided support for a reciprocal effects model: both self-efficacy and performance had statistically significant, unique positive influences on each other over time (Figure 2).

Interpretation. The finding of reciprocal effects is consistent with the conception of reciprocal determinism in social cognitive theory (Bandura, 1977, 1997; Pajares & Usher, 2008). Reciprocity aligns with notions of the self-efficacy ↔ performance relationship as a feedback loop, cycle or spiral (Klassen & Usher, 2010; Lindsley et al., 1995; Multon et al., 1991; Shea & Howell, 2000). The net performance → self-efficacy effect highlights the important role of performance experience in the formation of self-efficacy beliefs (Britner & Pajares, 2006; Usher & Pajares, 2008). The finding of a unique positive self-efficacy → performance influence while controlling for the opposite pathway in the CLPA casts doubt on the claim that self-efficacy is merely reflective of past performance (Bandura, 2012; Feltz et al., 2008; cf. Heggestad & Kanfer, 2005; Vancouver et al., 2001). Instead, this finding suggests that performance "is not simply a matter of how capable you are; it is also a matter of how capable you believe you are" (Pajares, 2006, p. 343). The unique influence of self-efficacy on performance highlights the generative or mobilising power of self-efficacy, which reflects *what you do with what you've got* (Gist & Mitchell, 1992). Consistent with this, a range of studies demonstrate that self-efficacy affects effort, resourcefulness, persistence, use of cognitive strategies, choice of activities, and goal setting (see Bandura, 1997, 2012; Klassen & Usher, 2010; Pajares, 1996; Pajares & Schunk, 2001).

In gauging the practical significance of these results, we note that the effect sizes for self-efficacy → performance across both models are in the range previously found by Valentine et al. (2004) and Richardson et al. (2012). The effect sizes for performance → self-efficacy are also similar in magnitude, though there are no studies, to our knowledge, with which we can directly compare these findings. On first observation, it may be argued that these effects, which, by convention are small-to-medium (Cohen, 1988), are unlikely to have far-reaching implications. However, academic achievement exists within a very intricate framework (Pajares, 2007). While self-efficacy is arguably one of the strongest non-intellective correlates of academic performance, many other factors are involved, both internal to the student; e.g., personality, coping styles (Komarraju, Karau, Schmeck, & Avdic, 2011; MacCann, Fogarty, Zeidner, & Roberts, 2011) and external to the student, e.g., parental engagement, student-teacher relationships (Dotterer & Lowe, 2011; Moreira, Dias, Vaz, &

Vaz, 2013). In such a complex context, the consistency and size of the present effects provide some evidence that they are indeed meaningful. The potency of the effect of self-efficacy on academic performance may have been overstated in the literature because of reliance on unidirectional findings. However, in the context of reciprocal effects, there is cause for optimism that interventions targeting either self-efficacy or performance will have flow-on effects to the other variable – and interventions which target both constructs will have synergistic effects. Conversely, considering either self-efficacy or performance in a vacuum potentially means ignoring half of the equation.

Research implications. The finding of a reciprocal self-efficacy ↔ performance relationship suggests that carefully timed measurement of both constructs will enhance research in this area. It is common in the literature to measure self-efficacy on multiple occasions followed by a single performance measurement, or to measure self-efficacy once, followed by staggered measures of performance. Researchers also measure self-variables on multiple occasions and then average the measurements. Such approaches are unlikely to provide accurate insights into how the two variables relate over time. Of course, research must be designed according to the specific research questions under consideration. Generally, however, researchers should consider reciprocity when attempting to isolate temporally dynamic features of the relationship (Bandura & Jourden, 1991).

Practical implications. The finding of a reciprocal self-efficacy ↔ performance relationship suggests that students would benefit when interventions to enhance self-efficacy are combined with regular opportunities to experience performance success. Approaches to enhancing academic self-efficacy are detailed elsewhere (Pajares, 2006; Siegle & McCoach, 2007). We provide several examples which are likely to have synergistic effects owing to their anticipated concurrent impact on self-efficacy and performance. Educators may observe that there is nothing novel here – what is suggested is that these approaches will be of most benefit when used in conjunction with each other in a cyclical fashion.

Guided mastery, which exists at the junction between self-efficacy and performance, involves assisting students to navigate through incremental authentic performance experiences, with scaffolding from teachers being reduced over time (Bong, 2012; Pajares & Schunk, 2001). It is recommended that tasks are graded to individual learners so that they are moderately challenging but accomplishable, and that teachers help students set incremental goals that are specific and proximal (Bong, 2012; Pajares, 2006). To provide evidence to students of their growing mastery, progress towards goals should be made explicit, feedback should be timely and accurate, and success attributed to internal, stable student characteristics/behaviours (Bong, 2012; Pajares, 2006; Siegle & McCoach, 2007).

An adaptive interpretation of performance experience may also leverage reciprocal self-efficacy ↔ performance effects. Performance should be framed in terms of progress rather than shortfalls, using internal rather than comparative standards, and showing how the experience can improve future performance (Bandura, 2012; Lindsley et al., 1995; Pajares, 2006). Where performance is interpreted as a failure, attributing this to factors under the student's control (e.g., insufficient effort or ineffective strategy use) can provide a buffer for self-efficacy beliefs (Bonne & Johnston, 2016; Pajares, 2006).

It must be noted that there is no single guaranteed way to enhance self-efficacy beliefs or performance outcomes, and it has been suggested that an assumption that it is always possible to generate lasting, adaptive changes to self-efficacy beliefs is neither accurate nor helpful (Gist & Mitchell, 1992; Klassen & Usher, 2010). Approaches should be explored with both contextual factors and the individual students in mind (Klassen & Usher, 2010).

Moderator effects: When believing leads to achieving, and vice versa

Age. As anticipated, the reciprocal relationship between self-efficacy and academic performance varied depending on participant age, with significantly stronger cross lagged effects for adults than for children. Importantly, the βSE1P2 cross-lagged path was not statistically significant for children, indicating that self-efficacy did not have a net effect on subsequent performance for children in this sample. That is, the relationship between self-efficacy and academic performance appears to be unidirectional for children, but reciprocal for adults. This suggests that, in children, self-efficacy may be detached from the generative processes generally ascribed to it, such as mobilisation of resources and strategies which lead to better performance, which is argued to exist in adults. This may be because the capacity to work with abstractions regarding the self is cultivated as cognitive maturity develops (Harter, 1999). Self-efficacy exists within a system of self-regulation, and individuals' self-regulatory capacities are known to change and develop over time (Bandura, 1981; Heckhausen & Dweck, 1998). Self-efficacy is influenced by a range of sources of information, in conjunction with a cognitive appraisal of that information (Bandura, 1997). It follows that, with greater cognitive maturity and more experience in educational settings, adults generate different and in particular more precise appraisals, for example, of previous performance outcomes or future task difficulty, with subsequent differences in self-efficacy levels (Phan, 2012; Phan & Ngu, 2016; Zimmerman, 1989). The differences between adults and children may also reflect a lack of calibration between self-efficacy and performance for children, which is also likely to develop over time and with an accumulation of task experience and feedback (Artino, 2012; Bandura, 2012).

Other differences between adults and children in each of the paths forming the overall model may also contribute; for example, self-efficacy appears to be more consistent over time for adults, which is consistent with the notion that self-efficacy stabilises as individuals approach the middle life span (Oettingen & Gollwitzer, 2015). On the other hand, performance is comparatively less stable over time for adults than children, likely due to the tendency for performance tasks to increase in difficulty more dramatically in adult educational settings (Pokay & Blumenfeld, 1990). In the same vein, it is also possible that age functions as a proxy for the type of academic tasks undertaken, with simpler and more defined tasks in childhood, and more complex and abstract tasks in adulthood.

One may be tempted to ascribe the lack of a significant finding for the β_{SE1P2} path to a lack of power to detect the effect, given that there are only two studies comprising children. However, there are several indications that this is unlikely to be the case. When considering the number of participants in each group, the sizes are roughly equal (adults, n=1182; children, n=1506). In addition, the observed power for the underlying pooled correlations is strong. Furthermore, we note that all other paths in the model comprising children are statistically significant, suggesting that the lack of a significant β_{SE1P2} path is a valid result.

These findings suggest that educators of adults should consider leveraging the synergies inherent in the self-efficacy ↔ performance relationship by using interventions which may enhance self-efficacy directly, while also providing regular opportunities for students to successfully perform academic tasks. Children may benefit from early opportunities to learn about and engage in metacognitive functions (Nietfeld, Cao, & Osborne, 2006), and to increase their understanding of, and exposure to, performance-feedback cycles (Artino, 2012; Gist & Mitchell, 1992; Gore Jr, 2006).

Methodological moderators: Lag, specificity match, self-efficacy scale type. It was found that studies that used relatively short time lags between first and second assessments (7 weeks or less) showed stronger cross-lagged relationships than studies with longer lag times (3-12 months), as expected. This is consistent with Bandura’s (1997) position regarding temporal disparity, which suggests that self-efficacy should be a more accurate predictor of behavioural variables over shorter time lags. In addition, it is likely that a range of unmeasured variables influence the paths in the model; it follows that there would be more intervening variables over a longer period of time, and also that these would impart a greater degree of influence over that period (Valentine et al., 2004). Considering the dearth of research regarding the impact of different time lags on the relationship between self-efficacy and performance, it is also possible to speculate that the longer time lags exceed the period during which the effects of interest are at their maximum levels (Clegg, Jackson, & Wall, 1977). In addition, we note that the reciprocal influences of self-efficacy and performance did not differ significantly from each other in studies conducted over a longer term. We speculate that over the course of repeated exposure to performance-feedback cycles, the dual influences of self-efficacy and performance may converge to a point of equilibrium.

Consistent with our hypothesis, stronger model paths were evident when the specificity of the self-efficacy and performance measures were matched than when they were not matched. This is consistent with both the theoretical position and recommendation for a micro-analytic research strategy (Bandura, 1977, 2012) and with empirical evidence in this regard (Bong, 2001; Pajares & Miller, 1995).

As anticipated, effect sizes were stronger for those studies which used unipolar scales than for those which used Likert scales. This finding is consistent with commentary above regarding the relative psychometric properties of different types of scales when measuring self-efficacy (Bandura, 2006, 2012; Klassen & Usher, 2010).

Overall, the moderator analyses found that measurement approaches congruent with those recommended by self-efficacy theorists (matched specificity, short lags and unipolar scales) resulted in stronger effects. While these differences did not consistently reach statistical significance, we note that the self-efficacy → performance effect sizes for those studies using theory-consistent approaches were consistently approximately twice the size of those found in the remaining studies. It appears that criticisms of self-efficacy measurement are going unheard by the research community, with negative ramifications for the quality of self-efficacy research (Klassen & Usher, 2010). Especially as longitudinal research exploring self-efficacy continues to develop, with increased need to detect more subtle changes over time, researchers are urged to make deliberate and well-informed decisions about the self-efficacy scales used, and to take heed of longstanding recommendations regarding self-efficacy measurement (Bandura, 2006; Bong, 2006; Lee & Bobko, 1994).

Sex and cohort effects. In this data set, sex was not identified as a significant moderator. While Huang (2013) found a gender difference in self-efficacy in a meta-analysis, the effect size was very small (g=0.08), and was calculated from 247 samples which were found to be highly heterogeneous, with individual effects ranging from g=-1.60 to g=1.40. Furthermore, findings of sex differences in self-beliefs and achievement are not consistent across meta-analyses, with Hansford and Hattie (1982) and Valentine et al. (2004) reporting no effect. It is possible that there was no such effect in the samples reviewed here, or that the number of included studies provided insufficient power to detect a very small effect.

Time since publication was not identified as a significant moderator, suggesting that cohort effects are not responsible for the patterns of variance observed in the included data.

Limitations

Given the narrow inclusion criteria of this study, a small number of studies was uncovered. This may be considered a limitation of the present analyses. However, we note that no formal minimum number of studies exists in order for meta-analysis to be useful – results from as few as two studies can be fruitfully synthesised (Valentine, Pigott, & Rothstein, 2010). Meta-analysis may not be suitable for a small group of studies showing wildly different effects. In this case, however, a range of assessments of the nature of the included data provide support for the validity of the present findings. For example, there are several indications that the correlation matrices used as input for the path analyses are stable: consistent effects across studies are shown in forest plots (Appendix 2.3), failsafe N values are substantial (Table 3), and there is strong underlying statistical power for the estimation of each of the model paths (Table 3). Furthermore, the lower bounds of the 95% confidence intervals for the pooled correlations not only confirm the statistical significance of the effects, but also suggest that the smallest possible estimates of effects based on the included data would still be deemed meaningful in the social sciences (Cohen, 1988). Nonetheless, the small number of studies does pose limitations in terms of interpretation: in particular, when interpreting the moderator analysis regarding age, we note that the studies focusing on children considered only the mathematical domain, and therefore we caution that this finding may not generalise to other areas of study. Given the paucity of studies uncovered for inclusion in this review, we echo the common call for more longitudinal research, particularly with children and culturally diverse participants.

Heterogeneity across studies was observed in the present review (Table 3), consistent with previous reports (Honicke & Broadbent, 2016; Multon et al., 1991; Richardson et al., 2012; Robbins et al., 2004; Valentine et al., 2004). However, the I^2 value of 45% for the self-efficacy → performance path for the SE-first studies (the path that is comparable to previous unidirectional meta-analyses) compares favourably to the values of ~70-90% reported in the only review which published comparable statistics (Richardson et al., 2012). By convention, an I^2 of 45% is indicative of moderate heterogeneity; values of 70-90% indicate substantial-to-considerable heterogeneity (Higgins & Green, 2008). Thus, we venture that the fine-grained approach taken has resulted in a more homogeneous group of studies for analysis. Even so, we note that the strict methodological inclusion criteria left little room to restrict the sample any further based on other contextual factors, such as domain of study, and that included studies may therefore vary in ways that were not addressed by our moderator analyses.

There is risk inherent in the statistical analysis in this study, owing to the fact that the pooled correlations used in the path analyses are subject to measurement error, whereas the path analysis assumes no such error (Selig & Little, 2012). However, given the stability of the input data as described above, this risk is mitigated.

Within self-efficacy theory, it is an individual's *interpretation* of their performance which is said to constitute mastery experience (Klassen & Usher, 2010; Usher & Pajares, 2008). For this reason, in research investigating the impact of mastery experience on self-efficacy, mastery experience is often operationalised with self-report questions such as *I got a good grade in science class last semester* (Britner & Pajares, 2006; Lent et al., 1991). This distinction becomes clear when considering an example: for a student with a C average, a B is likely to be interpreted as a success, whereas, for a straight-A student, a B may be deemed a failure (Pajares, 2006) – the B itself has no inherent value (Pajares & Usher, 2008). By using objective performance as a variable in this review, the interpretative mechanism which links objective performance to self-efficacy may be obscured in some cases.

With the search strategy we implemented, we were able to *exclude* studies which called their construct self-efficacy but measured it in a way inconsistent with social cognitive theory, but we were not able to *include* studies which measured a construct consistent with social cognitive theory but called something else, rendering our search potentially incomplete. Given that more than 5,000 studies were identified in the present search, an approach which would have ensured the inclusion of theoretically consistent but unlabelled self-efficacy studies seems unfeasible, particularly considering that researchers conducting studies within a social cognitive framework seem likely to use theory-consistent terminology.

The present review, being based on non-experimental data, precludes firm conclusions about causality. In this case, the concern is that the model is not fully specified, owing to the abundance of factors affecting human behaviour (Rogosa, 1988; Selig & Little, 2012). A finding of reciprocal effects may point to a blurring of the underlying processes which lead to self-efficacy and then to academic performance, or vice versa. Experiments conducted at a micro-analytic level are required when the aim is to detect causality, and further experimental research in academic settings is recommended. That being said, it has been noted that self-beliefs do not readily lend themselves to experimental manipulation: successful manipulation is likely to concurrently modify other, potentially unmeasured, mental processes (Valentine et al., 2004). In this context, cross-lagged panel analysis is considered to be “an important tool for building an argument for the causal effect of one variable on another” (Selig & Little, 2012, p. 271). In an area where complete experimental control is unlikely to be achieved, these types of analyses are of considerable value.

Directions for future research

While all included studies contained measures of both self-efficacy and performance on two occasions, in most cases (k=7), full correlation matrices were not published and were sought from the authors. Therefore, not only are there very few studies that collect panel data on these constructs, there are even fewer that leverage this data to explore reciprocity and the strength of directional effects. The lack of available information regarding SES of participants precluded a moderator analysis on that basis, and the lack of variability in terms of participant race and location meant that we were unable to undertake moderator analyses for these variables. Most participants were white North-Americans; this may affect the generalisability of our findings. These issues point to the need for further longitudinal research generally, and particularly with socially and culturally diverse participants.

Children were also under-represented in included studies. Given the contrasting findings for adults and children, the literature would be enhanced by an understanding of how the self-efficacy ↔ performance relationship changes over the course of childhood and adolescence. It appears that the relationship may change from unidirectional to reciprocal over this period. In the present analysis, where a gross distinction had to be made between adults and children, it was not possible to explore this.

We note that additional waves of self- data were available in some included studies, but could not be incorporated into the current analyses. Further meta-analytic research using change modelling approaches is warranted when sufficient primary studies become available.

Conclusion

In this systematic review with meta-analytic cross-lagged panel analysis, a reciprocal relationship between self-efficacy and academic performance was found. This finding is consistent with Bandura's view of self-efficacy within a framework of reciprocal determinism (Bandura, 1977, 1997) and suggests that interventions designed to influence either or both constructs may be fruitfully applied. To make a financial analogy, investment in either self-efficacy or performance improvements is likely to yield compound interest. In developing curricula designed to improve student performance, combining direct enhancement of self-efficacy with regular opportunities for performance success is likely to take best advantage of the identified reciprocal effects.

The self-efficacy ↔ performance relationship varied as a function of participant age, self-efficacy scale type, measurement lag and match of specificity between measures. A reciprocal relationship emerged for adults, but this did not hold for children. Across methodological moderators, those studies with approaches consistent with self-efficacy theory consistently showed stronger reciprocal effects.

In closing, we note that the effects reported here are embedded within an extremely complex system in which factors both internal and external to the student influence both self-efficacy and academic performance and the subtleties of their interrelation over time. As a consequence, no one approach can be recommended as a completely comprehensive or infallible method for improving academic performance or bolstering self-efficacy – the educational context and the individual learners must be taken into account.

References

References marked with an asterisk were included in the meta-analysis.

Artino, A. R. (2012). Academic self-efficacy: From educational theory to instructional practice. *Perspectives on Medical Education, 1*(2), 76-85. doi:10.1007/s40037-012-0012-5

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191. doi:10.1037/0033-295X.84.2.191

Bandura, A. (1981). Self-referent thought: A developmental analysis of self-efficacy. In J. H. Flavell & L. Ross (Eds.), *Social cognitive development: Frontiers and possible futures* (pp. 200-239). Cambridge: Cambridge University Press.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 5, pp. 307-337). Greenwich, CT: Information Age.

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management, 38*(1), 9-44. doi:10.1177/0149206311410606

Bandura, A. (2015). On deconstructing commentaries regarding alternative theories of self-regulation. *Journal of Management, 41*(4), 1025-1044.

Bandura, A., & Jourden, F. J. (1991). Self-regulatory mechanisms governing the impact of social comparison on complex decision making. *Journal of Personality and Social Psychology, 60*(6), 941-951.

Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology, 88*(1), 87-99. doi:10.1037/0021-9010.88.1.87

*Beck, J. W., & Schmidt, A. M. (2013). State-level goal orientations as mediators of the relationship between time pressure and performance: A longitudinal study. *Journal of Applied Psychology, 98*(2), 354-363. doi:10.1037/a0031145

Bong, M. (2001). Role of self-efficacy and task-value in predicting college students' course performance and future enrollment intentions. *Contemporary Educational Psychology, 26*(4), 553-570. doi:10.1006/ceps.2000.1048

Bong, M. (2006). Asking the right question: How confident are you that you could successfully perform these tasks? In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 287-305). Greenwich, CT: Information Age Publishing.

Bong, M. (2012). Self-efficacy. In J. A. C. Hattie & E. M. Anderman (Eds.), *International guide to student achievement* (pp. 64-66). New York: Routledge.

*Bonne, L., & Johnston, M. (2016). Students' beliefs about themselves as mathematics learners. *Thinking Skills and Creativity, 20*, 17-28. doi:10.1016/j.tsc.2016.02.001

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009a). Fixed effect vs. random effects models. *Introduction to meta-analysis* (pp. 77-86). Chichester, UK: John Wiley & Sons, Ltd.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009b). Meta-regression. *Introduction to meta-analysis* (pp. 187-203). Chichester, UK: John Wiley & Sons, Ltd.

Britner, S. L., & Pajares, F. (2006). Sources of science self-efficacy beliefs of middle school students. *Journal of Research in Science Teaching*, *43*(5), 485-499. doi:10.1002/tea.20131

Brown, S. D., Tramayne, S., Hoxha, D., Telander, K., Fan, X., & Lent, R. W. (2008). Social cognitive predictors of college students' academic performance and persistence: A meta-analytic path analysis. *Journal of Vocational Behavior, 72*(3), 298-308. doi:10.1016/j.jvb.2007.09.003

Caprara, G. V., Vecchione, M., Alessandri, G., Gerbino, M., & Barbaranelli, C. (2011). The contribution of personality traits and self-efficacy beliefs to academic achievement: A longitudinal study. *British Journal of Educational Psychology, 81*(1), 78-96. doi:10.1348/2044-8279.002004

Chiang, Y. T., & Lin, S. S. J. (2014). The measurement structure, stability and mediating effects of achievement goals in math with middle-school student data. *Scandinavian Journal of Educational Research, 58*(5), 513-527. doi:10.1080/00313831.2012.732607

Chin, D., & Kameoka, V. A. (2002). Psychosocial and contextual predictors of educational and occupational self-efficacy among Hispanic inner-city adolescents. *Hispanic Journal of Behavioral Sciences, 24*(4), 448-464. doi:10.1177/0739986302238214

Clegg, C., Jackson, P., & Wall, T. (1977). The potential of cross-lagged correlation analysis in field research. *Journal of Occupational Psychology, 50*(3), 177-196. doi:10.1111/j.2044-8325.1977.tb00374.x

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

*Daniels, M. A., Kain, J. M., Gillespie, J. Z., & Schmidt, A. M. (2010). *Latent profiles of goal orientation and motivational outcomes*. Paper presented at the 70th annual Academy of Management meeting, Montreal, Canada.

Davis-Kean, P. E., Huesmann, L. R., Jager, J., Collins, W. A., Bates, J. E., & Lansford, J. E. (2008). Changes in the relation of self-efficacy beliefs and behaviors across development. *Child development, 79*(5), 1257-1269. doi:10.1111/j.1467-8624.2008.01187.x

Dotterer, A., & Lowe, K. (2011). Classroom context, school engagement, and academic achievement in early adolescence. *Journal of Youth and Adolescence, 40*(12), 1649-1660. doi:10.1007/s10964-011-9647-5

Eccles, J. S., & Blumenfeld, P. (1985). Classroom experiences and student gender: Are there differences and do they matter? In L. C. Wilkinson & C. B. Marrett (Eds.), *Gender influences in classroom interaction* (pp. 79-113). New York: Academic.

Feltz, D. L. (1982). Path analysis of the causal elements in Bandura's theory of self-efficacy and an anxiety-based model of avoidance behavior. *Journal of Personality and Social Psychology, 42*(4), 764-781. doi:10.1037/0022-3514.42.4.764

Feltz, D. L., Chow, G. M., & Hepler, T. J. (2008). Path analysis of self-efficacy and diving performance revisited. *Journal of Sport and Exercise Psychology, 30*(3), 401-411. doi:10.1123/jsep.30.3.401

*Finney, S. J., & Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology, 28*(2), 161-186. doi:10.1016/S0361-476X(02)00015-2

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*(10), 906-911. doi:10.1037/0003-066X.34.10.906

Galla, B. M., Wood, J. J., Tsukayama, E., Har, K., Chiu, A. W., & Langer, D. A. (2014). A longitudinal multilevel model analysis of the within-person and between-person effect of effortful engagement and academic self-efficacy on academic performance. *Journal of School Psychology, 52*(3), 295-308. doi:10.1016/j.jsp.2014.04.001

*Galyon, C. E., Blondin, C. A., Yaw, J. S., Nalls, M. L., & Williams, R. L. (2012). The relationship of academic self-efficacy to class participation and exam performance. *Social Psychology of Education, 15*(2), 233-249. doi:10.1007/s11218-011-9175-x

Garriott, P. O., & Flores, L. Y. (2013). The role of social cognitive factors in Mexican American students' educational goals and performance: A longitudinal analysis. *Journal of Latina/o Psychology, 1*(2), 85-94. doi:10.1037/a0032378

Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management Review, 17*(2), 183-211. doi:10.5465/AMR.1992.4279530

Goldstein, H. (1983). Measuring changes in educational attainment over time: Problems and possibilities. *Journal of Educational Measurement, 20*(4), 369-377. doi:10.1111/j.1745-3984.1983.tb00214.x

Gore Jr, P. A. (2006). Academic self-efficacy as a predictor of college outcomes: Two incremental validity studies. *Journal of Career Assessment*, *14*(1), 92-115. doi:10.1177/1069072705281367

Hansford, B. C., & Hattie, J. A. (1982). The relationship between self and achievement/performance measures. *Review of Educational Research, 52*(1), 123-142. doi:10.3102/00346543052001123

Harter, S. (1999). *The construction of the self: A developmental perspective*. New York: Guilford Press.

Heckhausen, J., & Dweck, C. S. (1998). *Motivation and self-regulation across the life span*. Cambridge: Cambridge University Press.

Heggestad, E. D., & Kanfer, R. (2005). The predictive validity of self-efficacy in training performance: Little more than past performance. *Journal of Experimental Psychology: Applied, 11*(2), 84-97. doi:10.1037/1076-898X.11.2.84

Higgins, J. P., & Green, S. (2008). *Cochrane handbook for systematic reviews of interventions version 5*. The Cochrane Collaboration: The Cochrane Collaboration.

Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review, 17*, 63-84. doi:10.1016/j.edurev.2015.11.002

Hornstra, L., Van Der Veen, I., Peetsma, T., & Volman, M. (2013). Developments in motivation and achievement during primary school: A longitudinal study on group-specific differences. *Learning and Individual Differences, 23*(1), 195-204. doi:10.1016/j.lindif.2012.09.004

Huang, C. (2013). Gender differences in academic self-efficacy: A meta-analysis. *European Journal of Psychology of Education, 28*(1), 1-35. doi:10.1007/s10212-011-0097-y

Hwang, M. H., Choi, H. C., Lee, A., Culver, J. D., & Hutchison, B. (2016). The relationship between self-efficacy and academic achievement: A 5-year panel analysis. *Asia-Pacific Education Researcher, 25*(1), 89-98. doi:10.1007/s40299-015-0236-3

Klassen, R. M. (2004). A cross-cultural investigation of the efficacy beliefs of South Asian immigrant and Anglo Canadian nonimmigrant early adolescents. *Journal of Educational Psychology, 96*(4), 731-742. doi:10.1037/0022-0663.96.4.731

Klassen, R. M., & Usher, E. L. (2010) Self-efficacy in educational settings: Recent research and emerging directions. *Advances in Motivation and Achievement: Vol. 16 Part A* (pp. 1-33). Bingley, UK: Emerald Group.

Kock, N. (2014). Advanced mediating effects tests, multi-group analyses, and measurement model assessments in PLS-based SEM. *International Journal of e-Collaboration, 10*(1), 1-13. doi:10.4018/ijec.2014010101

Komarraju, M., Karau, S. J., Schmeck, R. R., & Avdic, A. (2011). The big five personality traits, learning styles, and academic achievement. *Personality and Individual Differences, 51*(4), 472-477. doi:10.1016/j.paid.2011.04.019

Lee, C., & Bobko, P. (1994). Self-efficacy beliefs: Comparison of five measures. *Journal of Applied Psychology, 79*(3), 364. doi:10.1037/0021-9010.79.3.364

Lee, H. S., Flores, L. Y., Navarro, R. L., & Kanagui-Munoz, M. (2015). A longitudinal test of social cognitive career theory's academic persistence model among Latino/a and white men and women engineering students. *Journal of Vocational Behavior*, *88*, 95-103. doi:10.1016/j.jvb.2015.02.003

Lent, R. W., Lopez, F. G., & Bieschke, K. J. (1991). Mathematics self-efficacy: Sources and relation to science-based career choice. *Journal of Counseling Psychology, 38*(4), 424-430.

*Lewis, J. L., Ream, R. K., Bocian, K. M., Cardullo, R. A., Hammond, K. A., & Fast, L. A. (2012). Con cariño: Teacher caring, math self-efficacy, and math achievement among Hispanic english learners. *Teachers College Record, 114*(7), 1-42.

Lindsley, D. H., Brass, D. J., & Thomas, J. B. (1995). Efficacy-performing spirals: A multilevel perspective. *Academy of Management Review, 20*(3), 645-678. doi:10.5465/amr.1995.9508080333

MacCann, C., Fogarty, G. J., Zeidner, M., & Roberts, R. D. (2011). Coping mediates the relationship between emotional intelligence (ei) and academic achievement. *Contemporary Educational Psychology, 36*(1), 60-70. doi:10.1016/j.cedpsych.2010.11.002

Majer, J. M. (2009). Self-efficacy and academic success among ethnically diverse first-generation community college students. *Journal of Diversity in Higher Education, 2*(4), 243-250. doi:10.1037/a0017852

Matsui, T., Matsui, K., & Ohnishi, R. (1990). Mechanisms underlying math self-efficacy learning of college students. *Journal of Vocational Behavior, 37*(2), 225-238. doi:10.1016/0001-8791(90)90042-Z

*Meier, S., McCarthy, P. R., & Schmeck, R. R. (1984). Validity of self-efficacy as a predictor of writing performance. *Cognitive Therapy and Research, 8*(2), 107-120. doi:10.1007/BF01173038

Menard, S. (2007). *Handbook of longitudinal research: Design, measurement, and analysis*. New York: Academic Press.

Mitchell, T. R., Hopper, H., Daniels, D., George-Falvy, J., & James, L. R. (1994). Predicting self-efficacy and performance during skill acquisition. *Journal of Applied Psychology, 79*(4), 506-517. doi:10.1037/0021-9010.79.4.506

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The prisma statement. *Annals of Internal Medicine, 151*(4), 264-269. doi:10.7326/0003-4819-151-4-200908180-00135

*Mone, M. A. (1994). Comparative validity of two measures of self-efficacy in predicting academic goals and performance. *Educational and Psychological Measurement, 54*(2), 516-529. doi:10.1177/0013164494054002026

*Mone, M. A., Baker, D. D., & Jeffries, F. (1995). Predictive validity and time dependency of self-efficacy, self-esteem, personal goals, and academic performance. *Educational and Psychological Measurement, 55*(5), 716-727. doi:10.1177/0013164495055005002

Moreira, P. A. S., Dias, P., Vaz, F. M., & Vaz, J. M. (2013). Predictors of academic performance and school engagement — integrating persistence, motivation and study skills perspectives using person-centered and variable-centered approaches. *Learning and Individual Differences, 24*, 117-125. doi:10.1016/j.lindif.2012.10.016

Moriarty, B. (2014). Research design and the predictive power of measures of self-efficacy. *Issues in Educational Research, 24*(1), 55-66.

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38*(1), 30-38. doi:10.1037/0022-0167.38.1.30

Muthen, L. K., & Muthen, B. O. (1998-2012). *Mplus user's guide. Seventh edition.* Los Angeles, CA: Muthen & Muthen.

Nietfeld, J. L., Cao, L., & Osborne, J. W. (2006). The effect of distributed monitoring exercises and feedback on performance, monitoring accuracy, and self-efficacy. *Metacognition and Learning, 1*(2), 159-179. doi:10.1007/s10409-006-9595-6

Nunnally, J. C., & Bernstein, I. H. (1994). The assessment of reliability. In I. H. Bernstein & J. C. Nunnally (Eds.), *Psychometric theory* (Vol. 3, pp. 248-292). New York: McGraw-Hill.

Obrentz, S. B. (2012). *Predictors of science success: The impact of motivation and learning strategies on college chemistry performance.* Georgia State University. Retrieved from http://scholarworks.gsu.edu/epse_diss/77

Oettingen, G., & Gollwitzer, P. M. (2015). *Self-regulation in adolescence*. Cambridge: Cambridge University Press.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543-578. doi:10.3102/00346543066004543

Pajares, F. (1997). Current directions in self-efficacy research. In M. L. Maehr & P. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 1-49). Greenwich, CT: JAI Press.

Pajares, F. (2006). Self-efficacy during childhood and adolescence: Implications for teachers and parents. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 339-368). Greenwich, CT: Information Age.

Pajares, F. (2007). Culturalizing educational psychology. In F. Salili & R. Hoosain (Eds.), *Culture, motivation, and learning: A multicultural perspective* (pp. 19-42). Charlotte, NC: IAP.

Pajares, F., Hartley, J., & Valiante, G. (2001). Response format in writing self-efficacy assessment: Greater discrimination increases prediction. *Measurement and evaluation in counseling and development, 33*(4), 214-221.

Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. *Contemporary Educational Psychology, 20*(4), 426-443. doi:10.1006/ceps.1995.1029

Pajares, F., & Miller, M. D. (1995). Mathematics self-efficacy and mathematics performances: The need for specificity of assessment. *Journal of Counseling Psychology, 42*(2), 190-198. doi:10.1037/0022-0167.42.2.190

Pajares, F., & Schunk, D. H. (2001). Self-beliefs and school success: Self-efficacy, self-concept, and school achievement. In R. Riding & S. Rayner (Eds.), *Perception* (pp. 239-266). London: Ablex Publishing.

Pajares, F., & Usher, E. L. (2008). Self-efficacy, motivation, and achievement in school from the perspective of reciprocal determinism. In M. L. Maehr, T. C. Urdan, & S. Karabenick (Eds.), *Advances in motivation and achievement* (Vol. 15, pp. 391-423).

Parker, P. D., Marsh, H. W., Ciarrochi, J., Marshall, S., & Abduljabbar, A. S. (2014). Juxtaposing math self-efficacy and self-concept as predictors of long-term achievement outcomes. *Educational Psychology, 34*(1), 29-48. doi:10.1080/01443410.2013.797339

Phan, H. P. (2012). Informational sources, self-efficacy and achievement: A temporally displaced approach. *Educational Psychology, 32*(6), 699-726. doi:10.1080/01443410.2012.708320

Phan, H. P., & Ngu, B. H. (2016). Sources of self-efficacy in academic contexts: A longitudinal perspective. *School Psychology Quarterly*, *31*(4), 548-564. doi:10.1037/spq0000151

Pokay, P., & Blumenfeld, P. C. (1990). Predicting achievement early and late in the semester: The role of motivation and use of learning strategies. *Journal of Educational Psychology, 82*(1), 41-50. doi:10.1037/0022-0663.82.1.41

*Richard, E. M., Diefendorff, J. M., & Martin, J. H. (2006). Revisiting the within-person self-efficacy and performance relation. *Human Performance, 19*(1), 67-87. doi:10.1207/s15327043hup1901_4

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353-387. doi:10.1037/a0026838

Riketta, M. (2008). The causal relation between job attitudes and performance: A meta-analysis of panel studies. *Journal of Applied Psychology, 93*(2), 472. doi:10.1037/0021-9010.93.2.472

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*(2), 261-288. doi:10.1037/0033-2909.130.2.261

Rogosa, D. (1988). Analysis of reciprocal effects. In J. P. Keeves (Ed.), *Educational research, methodology and measurement: An international handbook*. Cambridge, MA: Elsevier Science & Technology Books.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638. doi:10.1037/0033-2909.86.3.638

Schwarzer, R., & McAuley, E. (2016). The world is confounded: A comment on Williams and Rhodes (2016). *Health Psychology Review, 10*(2), 133-135. doi:10.1080/17437199.2016.1162667

Scott, L. A. (2000). *A matter of confidence? A new (old) perspective on sex differences in mathematics achievement.* Loyola University Chicago. PsycINFO database. (AAI9955419)

Selig, J. P., & Little, T. D. (2012). Autoregressive and cross-lagged panel analysis for longitudinal data. In B. Laursen, T. D. Little, & N. A. Card (Eds.), *Handbook of developmental research methods* (pp. 265-278). New York: Guilford Press.

Shea, C. M., & Howell, J. M. (2000). Efficacy-performance spirals: An empirical test. *Journal of Management, 26*(4), 791-812. doi:10.1177/014920630002600409

Siegle, D., & McCoach, D. B. (2007). Increasing student mathematics self-efficacy through teacher training. *Journal of Advanced Academics, 18*(2), 278-312. doi:10.4219/jaa-2007-353

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2015). Official user-guide to the p-curve. Retrieved from www.p-curve.com/guide.pdf

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2015). Better p-curves: Making p-curve analysis more robust to errors, fraud, and ambitious p-hacking, a reply to ulrich and miller (2015). *Journal of Experimental Psychology: General, 144*(6), 1146-1152. doi:10.1037/xge0000104

Sitzmann, T., & Yeo, G. (2013). A meta-analytic investigation of the within-person self-efficacy domain: Is self-efficacy a product of past performance or a driver of future performance? *Personnel Psychology, 66*(3), 531-568. doi:10.1111/peps.12035

Smith, M. L. (1980). Publication bias and meta-analysis. *Evaluation in Education, 4*, 22-24. doi:10.1016/0191-765X(80)90004-X

Stankov, L., & Lee, J. (2014). Quest for the best non-cognitive predictor of academic achievement. *Educational Psychology, 34*(1), 1-8. doi:10.1080/01443410.2013.858908

Thompson, S. G., & Sharp, S. J. (1999). Explaining heterogeneity in meta-analysis: A comparison of methods. *Statistics in medicine*, *18*(20), 2693-2708. doi:10.1002/(SICI)1097-0258(19991030)18:20<2693::AID-SIM235>3.0.CO;2-V

Tyagi, T. K., & Singh, B. (2014). The application of cross–lagged panel analysis in educational research. *Facta Universitatis: Philosophy, Sociology, Psychology and History, 13*(2), 39-51.

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research, 78*(4), 751-796. doi:10.3102/0034654308321456

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational psychologist, 39*(2), 111-133. doi:10.1207/s15326985ep3902_3

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics, 35*(2), 215-247. doi:10.3102/1076998609346961

*Vancouver, J. B., & Kendall, L. N. (2006). When self-efficacy negatively relates to motivation and performance in a learning context. *Journal of Applied Psychology, 91*(5), 1146-1153. doi:10.1037/0021-9010.91.5.1146

Vancouver, J. B., Thompson, C. M., & Williams, A. A. (2001). The changing signs in the relationships among self-efficacy, personal goals, and performance. *Journal of Applied Psychology*, *86*(4), 605. doi:10.1037/0021-9010.86.4.605

Viechtbauer, W. (2010). Conducting meta-analyses in with the metafor package. *Journal of Statistical Software*, *36*(3), 1-48. doi:10.18637/jss.v036.i03

Williams, T., & Williams, K. (2010). Self-efficacy and performance in mathematics: Reciprocal determinism in 33 nations. *Journal of Educational Psychology, 102*(2), 453-466. doi:10.1037/a0017271

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81*(3), 329. doi:10.1037/0022-0663.81.3.329

Zimmerman, B. J. (1995). Self-efficacy and educational development. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 202-231). Cambridge, UK: Cambridge University Press.

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25*(1), 82-91. doi:10.1006/ceps.1999.1016

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29*(3), 663-676. doi:10.2307/1163261

Zusho, A., Pintrich, P. R., & Coppola, B. (2003). Skill and will: The role of motivation and cognition in the learning of college chemistry. *International Journal of Science Education, 25*(9), 1081-1094. doi:10.1080/0950069032000052207

Chapter 3: Study 2

Reciprocity in the self-efficacy ↔ academic performance relationship:

The effect of measurement order

Abstract

Recent research addressing the chicken-and-egg conundrum regarding self-efficacy and academic performance shows that the two variables are reciprocally related. Using meta-analytic cross-lagged panel analysis (CLPA) of two-wave longitudinal studies, the present study builds on this foundation by exploring whether reciprocity holds when performance is measured prior to self-efficacy on each individual measurement occasion. Traditionally, self-efficacy is measured prior to performance ("SE-first"). However, exploring the effects of studies in which performance is measured prior to self-efficacy ("P-first") may shed light on the dynamics of the feedback loop which is believed to underlie the relationship. A preliminary analysis of SE-first (k=11, N=2,688) and P-first studies (k=3, N=769) indicated that the self-efficacy ↔ academic performance relationship was moderated by measurement order. Pooled correlations from the P-first studies were subsequently fitted to a CLPA model that provided further support for reciprocal effects. This analysis showed an opposite pattern in the strength of directional effects when compared with studies in which self-efficacy was measured prior to performance, with the unique impact of self-efficacy on performance (β=.149, p<.001) exceeding the net influence of performance on subsequent self-efficacy (β=.117, p<.001). We propose that this difference highlights the effect of task experience on the accuracy of self-efficacy beliefs, with self-efficacy and academic performance becoming more calibrated with each other over time.

Key words: *Self-efficacy, academic performance, reciprocal effects, calibration, meta-analysis*

Self-efficacy is a core dimension of human agency which refers to an individual's perception of their own capability to organise and execute required courses of action to achieve particular outcomes (Bandura, 1977, 1997). An established pedigree of reviews and meta-analyses identifies self-efficacy as a critical variable in determining academic success (e.g., Honicke & Broadbent, 2016; Klassen & Usher, 2010; Multon et al., 1991; Pajares & Schunk, 2001; Richardson et al., 2012). However, reliance on unidirectional research approaches has meant that the chicken-and-egg conundrum regarding the direction of causality in the relationship between self-efficacy and performance has only recently begun to be addressed (see Talsma, Schüz, Schwarzer, & Norris, 2017, for review).

Recent research findings provide support for the view that the relationship between self-efficacy and academic performance is reciprocal, with the two factors moderating each other iteratively in a feedback loop (Hwang, Choi, Lee, Culver, & Hutchison, 2016; Talsma et al., 2017). In a systematic review and meta-analytic cross-lagged path analysis of studies in which both self-efficacy and performance were measured on two occasions (see Figure 1), self-efficacy had a unique impact on subsequent performance (β_{SE1P2}=.071) and performance similarly uniquely impacted subsequent self-efficacy (β_{SE2P1}=.205) (Talsma et al., 2017). These findings are consistent with social cognitive theory, which embeds self-efficacy in a framework in which behaviour both shapes, and is shaped by a range of interacting factors (Bandura, 1977, 1997).

In their systematic review, Talsma and colleagues identified three studies which did not meet inclusion criteria because self-efficacy was measured subsequent to performance on each performance wave. While, strictly speaking, such an approach is not consistent with recommendations from self-efficacy theory to measure self-efficacy prior to performance (Moriarty, 2014), the fact that this practice occurs provides an opportunity to explore the potential effects of taking this approach. In particular, findings of studies measuring self-efficacy after performance may highlight different underlying mechanisms in the self-efficacy ↔ performance relationship, with the strength of directional effects affected by measurement order. If reciprocal effects are moderated by measurement order, this would have important implications for theory and research in the area, and may also impact on the development of educational policy regarding self-efficacy.

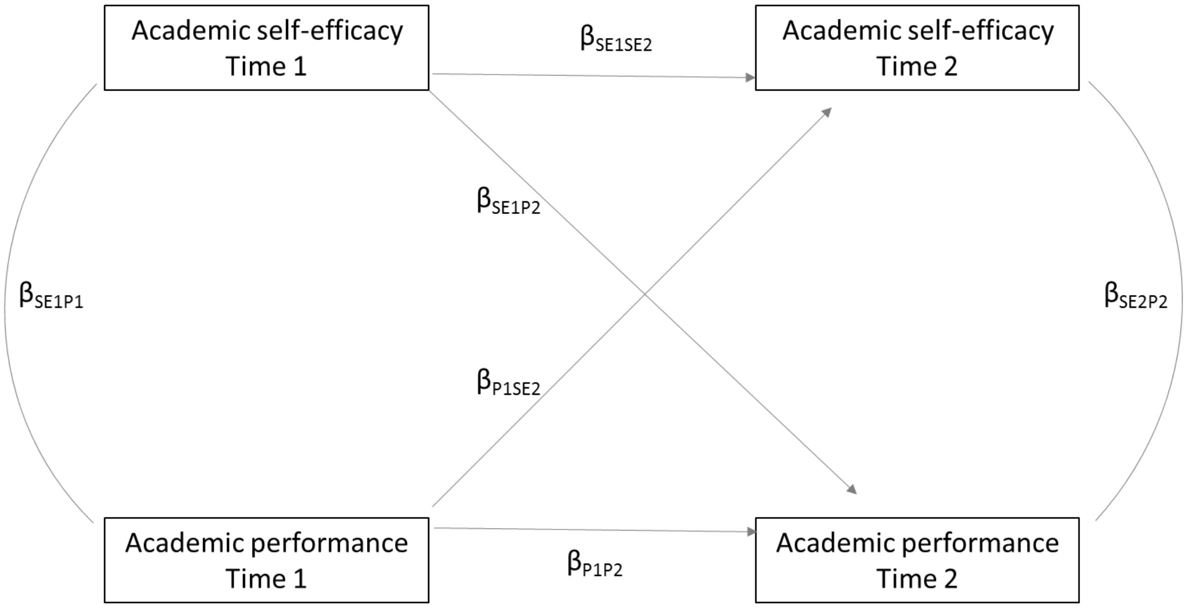


Figure 1. Conceptual diagram of the cross-lagged path model

When participants undertake a performance task before self-efficacy is measured, they gain mastery experience which is argued to be the principal source of information used in the formation of self-efficacy beliefs (Bandura, 1997). Gist and Mitchell's (1992) theoretical model of the self-efficacy/performance relationship provides a detailed framework for explaining how task experience may affect self-efficacy judgements. In their model, self-efficacy judgements are based on an analysis of task requirements (*what do I have to do?*), as well as an attributional analysis of experience (*how successfully have I done this before?*).

These two analyses together address the likelihood of successful future performance, and both are likely to be affected by direct task experience.

Individuals previously exposed to a specific performance task will have a clearer understanding of task requirements, and can draw directly on their experience to reflect on why a particular performance outcome occurred, leading to more accurate judgements of self-efficaciousness (Gist & Mitchell, 1992; Pajares & Schunk, 2001). Performance experiences which do not align with expectations are likely to affect subsequent self-efficacy beliefs by generating a self-corrective cycle in the self-efficacy/performance spiral (Gist & Mitchell, 1992; Shea & Howell, 2000) with continued performance experiences and feedback gradually reducing the gap between self-efficacy and performance (Lindsley et al., 1995; Pajares, 1997).

Conversely, for individuals who are unfamiliar with the task, ambiguity about its requirements may lead to inaccurate judgements of capacity to perform (Shea & Howell, 2000). Self-efficacy judgements made in this way may reflect an underestimation of the demands of the task, resulting in the tendency for individuals to overestimate their capacity to perform (Pajares & Kranzler, 1995). Self-efficacy judgements made prior to a performance task are also likely to be less accurate because individuals will need to either generalise their beliefs from more distal and perhaps vaguely recalled experiences which are perceived to be similar to the upcoming performance task, or rely more heavily on more general beliefs about ability (Chen, 2003; Gist & Mitchell, 1992; Mitchell, Hopper, Daniels, George-Falvy, & James, 1994; Pajares & Schunk, 2001).

In the context of this theoretical framework, studies in which self-efficacy is measured after participants have been exposed to the performance task may show a stronger relationship between self-efficacy on one occasion and performance on a subsequent occasion, owing to greater accuracy of self-efficacy beliefs which are informed by recent mastery experience.

Self-efficacy is a key construct in educational research, and research in this area is an important driver of educational policy (Pajares & Usher, 2008; Zimmerman, 1995). As such, further research is required to determine whether reciprocity and the strength of directional effects in the self-efficacy ↔ academic performance are invariant across different contexts.

The main aims of the present study were to explore reciprocity between self-efficacy and academic performance, and the comparative strength of net cross-lagged effects (β_{SE1P2} and β_{SE2P1}; see Figure 1), in those studies in which performance was measured prior to self-efficacy at each measurement wave; and to compare the strength of each of the cross-lagged paths estimated in the present study with its counterpart in the Talsma et al. meta-analysis.

Method

Inclusion criteria, literature search, and data extraction

We refer readers to the Talsma et al. meta-analysis for details regarding the rationale for inclusion criteria. In brief, studies were included if they:

1. Presented data regarding academic self-efficacy, measured in accordance with self-efficacy theory, at any level of specificity.
2. Collected data regarding academic performance on a specific occasion.
3. Collected data on self-efficacy and performance using a panel design in accordance with the path model shown in Figure 1, providing a set of six correlations.
4. Included participants in traditional primary, secondary and tertiary educational settings.
5. No interventions were conducted between the measurement waves, as interventions may obscure the effects of self-efficacy and academic performance on each other. Control group data were eligible for inclusion if other inclusion criteria were met.

Complete details regarding the literature search and coding are also to be found in the Talsma et al. (2017) meta-analysis. In brief, electronic searches were conducted in Scopus, PsycINFO, the Education Resources Information Centre (ERIC) and Web of Science. Three search fields covered terms relating to self-efficacy, academic performance (e.g., school, education, test), and methodological terms (e.g., panel, longitudinal, lag). Manual searches were also conducted using the reference lists of known extant meta-analyses which included analyses of the relationship between self-efficacy and academic performance.

Using the above criteria, 11 studies (N=2,688) were identified for inclusion in the original review. Three studies (N=769) which otherwise met all inclusion criteria were excluded on the basis that self-efficacy was measured subsequent to performance at each measurement wave, in an approach that was not considered to be consistent with the future-orientation of the self-efficacy construct (Schwarzer & McAuley, 2016). These three studies are of primary interest in the present meta-analysis.

Analysis

The CLPA model (Figure 1) comprises two auto-correlations (β_{SE1SE2} and β_{P1P2}), two cross-sectional correlations (β_{SE1P1} and β_{SE2P2}), and two cross-lagged correlations (β_{SE1P2} and β_{P1SE2}). To meta-analyse the findings of existing studies, six pooled correlations from included studies serve as input for the six paths in the model. Significant positive cross-lagged path coefficients demonstrate a reciprocal relationship, with the larger coefficient marking the stronger antecedent. Interested readers are directed to Talsma et al. (2017) for a detailed rationale regarding this methodological approach.

A preliminary moderator analysis was undertaken to ascertain whether measurement order significantly moderated any of the six pooled correlations. Using the 11 studies included in the Talsma et al. (2017) meta-analysis, along with the three additional studies with the reversed measurement order, a meta-regression analysis was conducted using Metafor for R (Viechtbauer, 2010). The cross-lagged path β_{SE1P2} (Figure 1) was significantly moderated by measurement order, Q=7.97, *p*=.005. This finding provided the impetus for exploring the main aims of the study.

Six pooled correlations were calculated from the zero-order correlations provided by the three P-first studies. Analyses were conducted on the zero-order correlations using *p*-curve 4.05 (Simonsohn, Nelson, & Simmons, 2015) to assess evidential value and calculate the underlying statistical power for the estimation of each of the six pooled correlations. Fisher's z transformations were conducted, and random effects modelling and restricted maximum likelihood estimation were used, with study weighting based on the inverse of study-specific variance. The matrix of pooled correlations resulting from the above analysis served as input for the CLPA (Figure 1) using MPLUS 7 (Muthen & Muthen, 1998-2012). Following procedures outlined by Kock (2014) using the pooled standard error method, the standardised betas representing the cross-lagged paths in this model were compared to the identical paths in the studies comprising the Talsma et al. (2017) meta-analysis.

Assessment of risk of bias

Moderator analyses regarding potential bias risk variables were not conducted, owing to the risk of invalid findings based on the small number of studies included. Risk of bias based on reliability of self-efficacy measures was considered to be low, given that all Cronbach's alphas for included studies exceeded recommended levels (Nunnally & Bernstein, 1994). It was not possible to assess risk of bias based on reliability of performance information as this was unavailable. However, data was taken from institutional records and is suggested to be at low risk of bias. Publication bias was assessed with multiple tests which are reported below. Inter-rater agreement on coding was 100% owing to the transparency of coding categories, suggesting a low risk of author bias.

Results

Features of included studies are shown in Table 1, while zero-order correlations from the included studies are presented in Table 2.

Table 1

Features of included studies

Study	n	Age group	% males	SE α	Lag	Specificity match	Scale
Chae, 2012	189	Adults	45	.95	Long	No	Unipolar
Kenney-Benson et al., 2006	518	Children	51	.85	Long	No	Likert
Schunemann et al., 2013	62	Children	43	.71	Long	Yes	Likert

Note: Short = ≤7 weeks; Long = ≥3 months; Specificity match = self-efficacy and performance measured at the same level of specificity (task/domain); see method section of Chapter 2 for detailed descriptions of categories.

Table 2

Zero-order correlations from included studies

Study	n	SE1P1	SE1P2	SE2P1	SE2P2	SE1SE2	P1P2
Chae, 2012	189	0.585	0.444	0.355	0.382	0.574	0.503
Kenney-Benson et al., 2006	518	0.440	0.290	0.290	0.470	0.370	0.560
Schunemann et al., 2013	62	0.185	0.388	0.086	0.346	0.473	0.395

Note: SE1P1 = self-efficacy at time 1 with performance at time 1; SE1P2 = self-efficacy at time 1 with performance at time 2; SE2P1 = self-efficacy at time 2 with performance at time 1; SE2P2 = self-efficacy at time 2 with performance at time 2; SE1SE2 = self-efficacy at time 1 with self-efficacy at time 2; P1P2 = performance at time 1 with performance at time 2. ** p<.01 ***p<.001

Pooled estimates of the six correlations in the cross-lagged panel model are shown in Table 3, along with standard errors, and 95% confidence intervals (forest and funnel plots are available in Appendices 3.1 and 3.2). Moderate-to-strong positive auto-correlations, and small-to-moderate positive synchronous and cross-lagged correlations were found, consistent with previous findings. Also shown in Table 3 are I^2 and Q heterogeneity test values; these figures suggest sampling error alone is unlikely to account for the differences in effect sizes, however, we note that the zero/near-zero values may not be reliable because of the small number of studies included (Kontopantelis, Springate, & Reeves, 2013). The results of the p-curve analyses also in Table 3 can be used to make inferences about evidential value (Simonsohn et al., 2015). A left-skewed distribution suggests that a data set is subject to selective reporting or *p*-hacking. In the present case, the full- and half- *p*-curve analyses all show right-skewed distributions, indicating that the data included in the calculation of each of the six model paths have evidential value. P-curve analysis also provides observed power estimates, which are reported in Table 3.

Figure 2 displays the meta-analytic path model estimated from the pooled correlations in Table 3. The results are consistent with a reciprocal effects model, with all paths positive and statistically significant, $p<.001$. The net self-efficacy → performance effect is $\beta=.149$, whereas the net performance → self-efficacy effect is $\beta=.117$. These estimates are based on a fully saturated model. A model was also tested in which the cross-lagged paths were constrained to be equal, indicating an equivalent reciprocal relationship (Berrington, Smith, & Sturgis, 2006; Menard, 2007). In this case, the constrained model did not provide a significantly poorer fit to the data ($\chi^2 = 0.463$, $p>.05$). Therefore, the influence of self-efficacy on performance cannot be said to be significantly greater than the reverse, though it is likely that the small sample size affected the calculation of the Chi-squared statistic in this case.

Table 3

Pooled correlations and tests of heterogeneity and publication bias

	k	n	SE1P1	SE1P2	SE2P1	SE2P2	SE1SE2	P1P2
Pooled correlation (SE)	3	769	.429*** (0.13)	.412*** (0.03)	.315*** (0.03)	.406*** (0.03)	.511*** (0.09)	.677*** (0.19)
95%CI			.167, .691	.341, .483	.244, .386	.335, .477	.337, .685	.314, 1.04
I^2			89.33	0	0.01	0	75.43	94.49
Q			14.98***	1.08	3.69	0.14	9.77**	42.27***
Failsafe N			136	124	60	112	171	423
p-curve, full (Z)			-10.97***	-8.03***	-7.4***	-7.96***	-10.66***	-9.6***
p-curve, half (Z)			-10.84***	-7.71***	-7.21***	-7.57***	-10.43***	-9.31***
Observed power			99%	99%	99%	99%	99%	99%
[90%CI]			[99%, 99%]	[99%, 99%]	[99%, 99%]	[99%, 99%]	[99%, 99%]	[99%, 99%]

Note: Correlations are Fisher's Z transformations. SE1P1 = self-efficacy at time 1 with performance at time 1; SE1P2 = self-efficacy at time 1 with performance at time 2; SE2P1 = self-efficacy at time 2 with performance at time 1; SE2P2 = self-efficacy at time 2 with performance at time 2; SE1SE2 = self-efficacy at time 1 with self-efficacy at time 2; P1P2 = performance at time 1 with performance at time 2. ** p<.01 ***p<.001

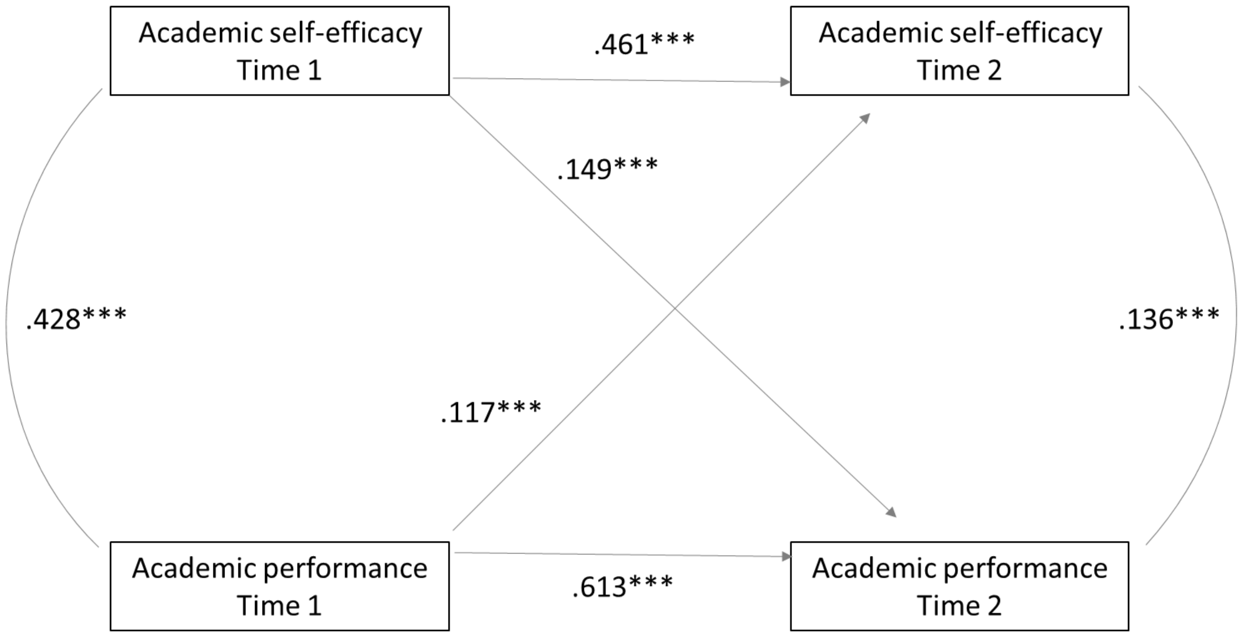


Figure 2. Cross-lagged panel model

Note: Figures are standardised path coefficients. A model in which the cross-lagged paths were constrained to be equal did not provide a significantly poorer fit to the data. ****p*<.001

The β_{SE1P2} cross-lagged path in the present model was significantly larger (*t*=2.32, *p*=.02) than the equivalent path in the model presented by Talsma et al. (2017). Similarly, the β_{P1SE2} cross-lagged path in the present model was significantly smaller (*t*=-2.76, *p*<.01) than the analogous path in the model analysing studies which measured self-efficacy prior to performance.

Discussion

This study further explored the chicken-and-egg conundrum in the relationship between academic self-efficacy and academic performance, meta-analysing studies which measured self-efficacy subsequent to performance on two separate occasions to consider the effect of measurement order on the self-efficacy ↔ performance relationship. The CLPA provided further support for a reciprocal effects model, consistent with Talsma et al. (2017): both self-efficacy and performance had statistically significant, unique positive paths to each other over time (see Figure 2). This finding is consistent with reciprocal determinism of self-efficacy and performance (Bandura, 1977, 1997; Pajares & Usher, 2008), and demonstrates the important effects of both self-efficacy (as a mobilising factor) and performance (as a source of mastery experience) in educational settings (Talsma et al., 2017). Even when there are methodological differences, reciprocity in the relationship appears to hold. The applied and theoretical implications of this reciprocity are discussed in further detail in Talsma et al. (2017); in the present discussion we focus on the findings regarding the strength of cross-lagged effects.

Compared with the Talsma et al. (2017) meta-analysis, the strength of directional effects was reversed: self-efficacy was a stronger predictor of performance (β=.149) than was performance of self-efficacy (β=.117). While the difference between the two standardised betas in the same model was not statistically significant (thus we cannot rule out an equivalent reciprocal relationship with the present data), the self-efficacy → performance path in the present model was significantly stronger than in the identical path in the model where self-efficacy was measured prior to performance. Conversely, the performance → self-efficacy path was significantly weaker than the analogous path in the Talsma et al. (2017) meta-analysis. The fact that the magnitude of the regression weights was the reverse of that found when self-efficacy was measured prior to performance suggests an important effect of measurement order.

Interpretation

These findings lend weight to the theoretical model outlined by Gist and Mitchell (1992) and described above. Specifically, we argue that participants in the present analysis were able to draw on their experience with the performance task to make more informed task and attributional analyses when forming their self-efficacy beliefs for future performance. We speculate that performance experience thus influenced the accuracy, or *calibration*, of self-efficacy beliefs; that is, the level of agreement, or otherwise, between self-efficacy judgements and actual performance (Artino, 2012; Stone, 2000). It may be argued that the observed pattern of cross-lagged effects reflects better calibration for those whose initial self-efficacy was measured following initial performance, as in the present study, and poorer calibration for those whose initial self-efficacy was measured prior to initial performance, as in Talsma et al. (2017), (also see Stone, 2000, for discussion). We posit that this difference in calibration is likely owing to higher levels of task experience, and therefore lower levels of task ambiguity, for individuals who engaged in the performance task prior to judging self-efficacy (Bandura, 2012; Gist & Mitchell, 1992; Lindsley et al., 1995; Stone, 2000). Simply put, we believe that our results reflect the fact that individuals who have had specific experience of a performance task are able to more accurately assess their ability to perform similar tasks in the future than individuals who have not had that specific experience.

Practical implications

Calibration of self-efficacy beliefs with academic performance is considered to be a desirable outcome: in order to be of most benefit to performance, self-efficacy beliefs should be reasonably accurate – marked over-confidence can lead to complacency, whereas substantial under-confidence can lead to disengagement (Artino, 2012; Bandura, 2012; Pajares, 2006; Stone, 2000). If increased accuracy of self-efficacy beliefs is indeed what underscores the difference in strength of cross-lagged effects in these models, as we argue, then some clues are provided as to how to enhance calibration. Most specifically, it appears that calibration may be enhanced by experience with the performance task. We draw on related research to build on this basic observation, noting that calibration is likely to be stimulated by early access to opportunities to perform and perpetuated by a steady succession of performance-feedback cycles (Artino, 2012; Mitchell et al., 1994; Pajares, 2006; Sitkin, 1992; Stone, 2000). However, research regarding approaches to improving calibration has yielded mixed results, and further research is needed (e.g., Bol & Hacker, 2012; Hacker, Bol, & Bahbahani, 2008; Nietfeld, Cao, & Osborne, 2006).

Research implications

The suppositions above regarding measurement order, the strength of directional effects, and calibration have some implications for research and theory. In particular, attempts to directly compare effects from studies which differ in order of measurement of self-efficacy and performance may be an exercise in comparing apples and oranges. To highlight this, consider the effect on our findings if studies using both orders of measurement had been included in a single meta-analysis: the opposite pattern in the strength of cross-lagged effects might have resulted in the directional effects washing each other out. Explicit reporting of order of measurement was lacking in almost half of the 14 studies referred to in this review, with such information only available on contacting authors. Given that differences in measurement order may be associated with different strengths of cross-lagged effects, it is critical that this information is published. The fact that both self-efficacy and performance were measured “in the same session” is insufficient.

Finally, we note that the self-efficacy ↔ performance relationship is likely to reflect a backward infinite regress, rendering any data collection over a finite period inescapably truncated. This challenge seems empirically insurmountable, without descending into nonsensicality: starting longitudinal research by asking babies how confident they are to perform grasping motions. Nonetheless, research may be improved by careful consideration of how best to design studies to take account of the impact of measurement order and timing, and how to control for covarying common causes. Decisions should be guided by a clear rationale, rather than the apparent trend of letting pragmatics be the primary consideration.

Limitations and directions for future research

Limitations of the present study and directions for future research are analogous to those outlined in detail in Talsma et al .(2017). In brief, two key concerns are an inability to make firm causal inferences, and small sample size. If the tested model does not contain all relevant variables, which in such a complex relationship is unlikely, then strong conclusions cannot be drawn regarding causality (Selig & Little, 2012). However, CLPA helps to build an argument for causality, and is especially informative in complex cases where the required degree of experimental control may be impossible to achieve (Berrington et al., 2006; Menard, 2007).

In the present analysis, readers may be particularly concerned about the small number of studies included. As was pointed out in Talsma et al. (2017) there is no specific minimum number of studies which can fruitfully be meta-analysed. The findings of these few studies are very consistent (as reflected in consistent forest plots and strong observed power, and in the relatively narrow CIs for the pooled correlations given such a small number of studies) giving cause for optimism that the results reported here are legitimate.

That being said, further research is certainly needed with regard to the effect of measurement order on the reciprocity and strength of directional effects in the self-efficacy ↔ performance relationship. With regard to calibration of self-efficacy and academic performance specifically, we note that research in this area is limited (cf., Chen & Zimmerman, 2007; Chen, 2003). Further research considering the accuracy of self-efficacy beliefs with respect to academic performance outcomes is warranted.

Conclusion

The present meta-analysis provides further evidence of reciprocal determinism in the relationship between self-efficacy and academic performance (Bandura, 1978; Pajares & Usher, 2008). Educators and education policy makers may anticipate small but meaningful positive effects on performance from increases in self-efficacy, and on self-efficacy from exposure to performance experience.

The finding of a different pattern in the strength of cross-lagged effects depending on the measurement order of the two key variables at each measurement wave is argued to be consistent with a calibration effect in the relationship, such that task experience enables individuals to more accurately form self-efficacy beliefs about future performance capacity. Such a finding, if borne out in future research, suggests how a steady performance-feedback cycle may facilitate calibration of self-efficacy beliefs with actual performance capacity.

Reciprocity and calibration in the self-efficacy ↔ performance relationship both provide promising avenues for future investigation. Future research is likely to be strengthened by designs which take into consideration the reciprocal nature of the relationship between self-efficacy and academic performance, as well as the issue of measurement order which appears to differentially affect the pattern of relationships between these variables. Importantly, researchers should be aware that studies approaching self-efficacy and performance from a uni-directional standpoint are likely neglecting a key facet of the dynamic of the relationship.

Chapter 4: Study 3

Miscalibration of self-efficacy and academic performance:

Self-efficacy ≠ self-fulfilling prophecy

Abstract

Whether academic self-efficacy aligns with performance capacity – and whether it *should* align – is a contested issue. Little is known about the accuracy of university students' self-efficacy beliefs for their course assessments, or how self-efficacy accuracy relates to future academic performance. University students (n=207) completed a series of self-efficacy questionnaires. Self-efficacy was compared to performance on two written assignments, two exams, and subject grades. Self-efficacy accuracy (the absolute difference between self-efficacy and performance) and bias (the signed difference; over- and under-efficaciousness) were used to predict subsequent performance. Miscalibration was prevalent. Under-efficaciousness was common for written assignments/exams; over-efficaciousness was pronounced for subject grades. Self-efficacy exceeded performance for low-achievers, but fell short of performance for higher achievers. Under-efficaciousness at T1 predicted stronger performance on similar tasks at T2 (written assignment, $\beta=-.337$; exam, $\beta=-.402$). Findings suggest self-efficacy is not a self-fulfilling prophecy. Miscalibrated students may be subject to negative impacts on self-regulation and performance.

Key words: *Self-efficacy, academic performance, accuracy, bias, calibration*

Calibration – the accuracy of subjective beliefs with regard to objective outcomes (e.g., Dunlosky & Thiede, 2013; Hacker, Bol, & Keener, 2008) – is an important issue in educational settings, as it is believed to influence academic behaviours and outcomes (e.g., Dunlosky & Rawson, 2012; Ramdass & Zimmerman, 2008; Zabrucky, 2010). Where self-evaluations exceed objective reality, students may cease studying prematurely, may be less inclined to seek help and support, and may exhibit complacency or carelessness (Boekaerts & Rozendaal, 2010; Hacker, Bol, Horgan, & Rakow, 2000; Zvacek, de Fátima Chouzal, & Restivo, 2015). Conversely, individuals who sell themselves short may waste time on overstudying material that is already known, uncritically adopt unhelpful suggestions from others, or disengage because of a sense of self-doubt (Boekaerts & Rozendaal, 2010; Hacker et al., 2000; Usher, 2016).

Most calibration research focuses on metacognitive judgements (Schraw, 2009), text comprehension (Zabrucky, 2010) and prediction of test performance (Hacker et al., 2000), thus asking whether students *know what they know*. However, given the pivotal role of self-beliefs in academic self-regulation (Zabrucky, 2010; Zimmerman, 2002), it is also important to ask: do students know what they are capable of achieving in academic settings? This is fundamentally a question of calibration of academic self-efficacy. Self-efficacy – perceptions of one's capability to organise and execute required courses of action to achieve particular outcomes (Bandura, 1997) – is widely believed to be one of the most important non-intellective determinants of academic performance (Richardson, Abraham, & Bond, 2012). Consistent significant positive correlations between self-efficacy and academic performance are often interpreted to mean that self-efficacy beliefs are accurate with respect to performance outcomes (e.g., Ackerman, Beier, & Bowen, 2002; Moores, Chang, & Smith, 2006), providing support for the idea of self-efficacy as a self-fulfilling prophecy – believe, and you will achieve (Pajares, 2006). However, even a large positive correlation may be underscored by consistent bias (Dunning & Helzer, 2014), and some research suggests a lack of correspondence between self-efficacy and academic performance (e.g., Chen & Zimmerman, 2007; Chen, 2003).

Are students' academic self-efficacy beliefs calibrated or biased?

Our first research question concerns calibration of students' academic self-efficacy beliefs. Previous studies in this area typically focus on mathematics self-efficacy and performance in children and adolescents; in these studies, self-efficacy beliefs have generally been shown to exceed capacity to perform – that is, students are over-efficacious (e.g., Chen & Zimmerman, 2007; Chen, 2003; Pajares & Graham, 1999; Pajares & Kranzler, 1995; Ramdass & Zimmerman, 2008). However, several factors limit the generalisability of the previous self-efficacy calibration research. A key issue is that this research has tended to focus on school-aged students, who are known to differ from adults in terms of the strength of the relationship between self-efficacy and academic performance, which suggests that more research with adult learners (e.g., university students) is necessary (e.g., Richardson, Abraham, & Bond, 2012).

Studies focusing on mathematics self-efficacy and performance also involve very particular methods which are not applicable in many other natural learning tasks. In the studies cited above, a typical approach is to ask participants to look at mathematics problems for a short period of time, and then rate their degree of confidence in correctly solving the problem. Immediately afterwards, participants solve the problems they were previously shown, or virtually identical problems. While this method provides a high degree of control and emulates the circumstances under which self-efficacy is known to be most predictive of learning outcomes (Bandura, 1997), there is some doubt as to whether such an approach is reflective of authentic learning experiences (Pieschl, 2009).

When considering the university experience, written assignments and exams are common assessment methods, and these differ from previously studied judgements about mathematics capacity in a number of ways which may influence calibration findings. Participants have a genuine stake in the outcomes of these types of tasks, which require a more complex range of self-regulated behaviours, are influenced by different degrees of motivation, and are subject to different timing dynamics and greater uncertainty than tasks designed for the purposes of an experiment and administered in a single sitting (Gist & Mitchell, 1992; Hacker et al., 2008; Mengelkamp & Bannert, 2009). In addition, studies undertaken regarding mathematics performance use localised judgements, which are made in relation to individual items (e.g., mathematics problems) and then averaged, whereas judgements referring to a whole task (e.g., an essay or exam) are more likely to occur naturally in educational settings (Pieschl, 2009). A further effect of the focus on isolated mathematics tasks in the existing literature is that little is known about calibration of self-efficacy beliefs with performance across a whole course of study.

We also note that some common lines of research in the calibration paradigm more broadly are underdeveloped in research regarding calibration of self-efficacy beliefs specifically, such as the relative prevalence of over- and under-estimations of ability (Stankov & Lee, 2014) and the tendency for stronger performers to underestimate their capacity, while weaker performers overestimate theirs (Bol, Hacker, O'Shea, & Allen, 2005; Hacker et al., 2008).

In sum, exploration of the self-efficacy calibration of university students undertaking authentic performance assessments is needed, to determine whether patterns of calibration are analogous to those identified in the previous literature. We turn now to our second key research question.

What level of self-efficacy is optimal with regard to academic performance outcomes: over-efficaciousness, under-efficaciousness, or accuracy of self-efficacy beliefs?

As the optimal level of self-efficacy in educational settings is a matter of discussion (e.g., Bandura & Locke, 2003; Vancouver & Kendall, 2006), it is also important to investigate how calibration of self-efficacy relates to future academic performance. Some researchers argue that self-efficacy beliefs which accurately reflect actual performance capacity will be of most benefit to performance (Stankov & Lee, 2017). Students who accurately judge what they can and cannot do are argued to be better able to adapt strategies effectively to the demands of a given task (Boekaerts & Rozendaal, 2010). Chen (2003) reported that greater accuracy of school children’s mathematics self-efficacy positively predicted subsequent mathematics performance. However, researchers are far from unanimous in this regard.

In a contrasting position, it is suggested that self-efficacy beliefs that exceed current capacity to perform are adaptive, motivating students to mobilise resources to increase performance above previous levels (Bandura, 1997). In this view, overestimation improves effort and persistence, and attempts to make students more realistic about their performance capacity is considered to be a dangerous enterprise (Pajares, 2006). Pajares argues that a reach that exceeds one’s grasp should be encouraged – because one’s sense of self-efficacy creates a self-fulfilling prophecy with respect to performance outcomes.

A third perspective is based on concerns that over-efficaciousness is instead potentially associated with a complacent attitude, whereby students may be content to coast along in their studies without exerting effort or appropriately monitoring their performance (e.g., Dunlosky & Rawson, 2012; Zabrucky, 2010). Support for this is found in discrepancy-reduction theories of self-regulation (Carver & Scheier, 1982); researchers argue that overly strong self-efficacy beliefs may obscure any discrepancy between the current and desired state of learning, leading to reduced effort and poorer performance (Vancouver & Kendall, 2006). While collinearity precluded the use of self-efficacy bias in Chen's (2003) regression model predicting subsequent mathematics performance, she noted that self-efficacy bias and subsequent performance were negatively correlated – such that under-efficaciousness was associated with stronger subsequent performance.

Much of the commentary on what constitutes the optimal level of self-efficacy amounts to conjecture, because very few studies have directly addressed how calibration on one occasion relates to future behaviour (Bol & Hacker, 2012; Dunlosky & Rawson, 2012). Thus, the second key aim of the present study is to analyse whether self-efficacy calibration on one occasion predicts subsequent performance outcomes, using data collected on two occasions with regard to exams and written reports. The three positions outlined above would lead to three different hypotheses: that better future performance could be predicted by either (1) accurate self-efficacy beliefs, (2) self-efficacy exceeding current performance (over-efficaciousness), or (3) self-efficacy underestimating current performance (under-efficaciousness). Given that there are plausible arguments for all three positions, the analyses regarding the prediction of performance from self-efficacy bias will be exploratory in nature.

Summary

To build on the extant literature outlined above, the present study aims to explore self-efficacy calibration of university students with regard to a range of authentic academic tasks, and to analyse how self-efficacy calibration relates to future academic performance. Firstly, we gauge the accuracy and bias (see measures in the method section) of students' self-efficacy beliefs by comparing self-efficacy judgements for five academic performance outcomes with the actual outcomes: two written reports, two multiple-choice exams, and overall subject grades. Following previous findings in the literature, we anticipate students will be inaccurate, and tend towards over-efficaciousness. We also posit that stronger students will be under-efficacious while weaker students will be over-efficacious. Secondly, by analysing two waves of data each for written reports and exams, we assess how self-efficacy calibration (accuracy and bias) at time 1 relates to future performance on the same type of task at time 2. The findings of these exploratory analyses will shed light on the debate regarding whether over-efficaciousness, under-efficaciousness, or accuracy of self-efficacy beliefs best predicts future academic performance outcomes.

Method

Participants and procedure

Participants were 207 first-year undergraduate psychology students (152 female, mean age 25, age range 18–66 years) from an Australian university, who received course credit for participation. The study design is illustrated in Figure 1. Data were collected via online questionnaires and from institutional records over one semester plus the subsequent examination period in 2014 and 2015. After reading the study information and providing informed consent, participants completed a baseline questionnaire regarding self-efficacy for overall subject grade performance. Four identical questionnaires were completed within the five-day periods leading up to the submission of two written reports and the completion of two exams. Five sets of corresponding grades were recorded. Ethics approval was obtained from the Tasmanian Social Sciences Human Research Ethics Committee.

Measures

Academic performance. Academic performance was measured following Australian standards, as the grade achieved for the two written reports, two exams, and subject overall, with a possible range of 0 to 4 (fail [<50%], pass [50-59%], credit [60-69%], distinction [70-79%], and high distinction [≥80%], respectively). To ensure reliability, institutional policy provides that assessments are marked using rubrics and are moderated for consistency.

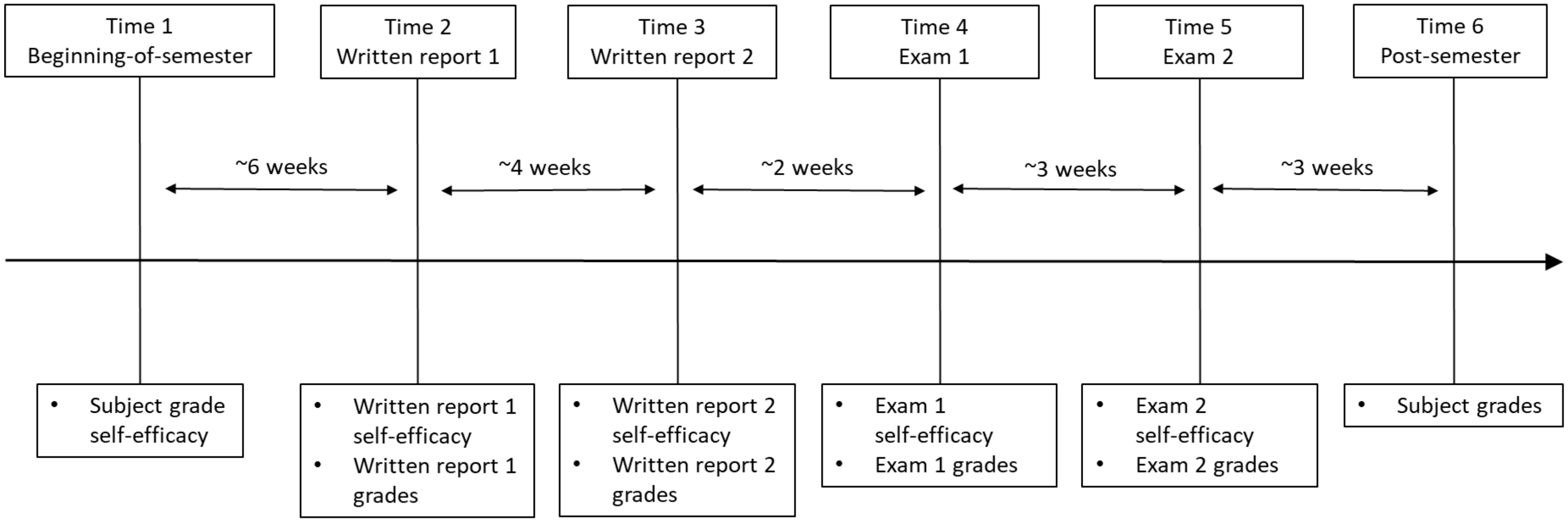


Figure 1. Longitudinal study design

Self-efficacy. Following Bandura's (2006) recommendations, in the online questionnaire, participants were presented with each of the performance levels (grades) in order of increasing difficulty and asked whether they were confident in their ability to achieve each one (yes/no; self-efficacy magnitude). Next, participants stated their level of confidence that they could attain each performance level (0–100%; self-efficacy strength). Following recommendations regarding the calculation of composite self-efficacy scores (Lee & Bobko, 1994), strength values (as decimals) were summed for each performance level that the participant indicated "yes" at magnitude level. E.g., a participant who responded "yes" they were confident in their ability to achieve a pass (with 100% confidence) and "yes" they were confident in their ability to achieve a credit (with 70% confidence), but "no" they were not confident in their ability to achieve a distinction (confidence 20%) would receive a score of 1.7 on the scale. Participants responding "no" to all binary decisions received a score of zero. Self-efficacy was measured on the same scale as performance, to facilitate the calculation of bias and accuracy scores.

Self-efficacy calibration. The three potential hypotheses regarding prediction of future performance from self-efficacy calibration require measures which allow us to determine the effects of under-efficaciousness, over-efficaciousness, and accurate self-efficacy beliefs. Consequently, both bias and accuracy measures were calculated.

Self efficacy bias. Self-efficacy bias is the signed deviation of the self-efficacy score from the academic achievement band score for each of the five performance outcomes (Dunlosky & Thiede, 2013). A bias score of zero reflects perfect calibration, positive values indicate over-efficaciousness, and negative values indicate under-efficaciousness.

Self-efficacy accuracy. Self-efficacy accuracy reflects only the magnitude of the deviation between self-efficacy and performance, ignoring direction. All values are positive, with scores closer to zero reflecting greater accuracy.

Analyses

To gauge overall accuracy, one-sample *t*-tests compared self-efficacy accuracy to a test value of zero (perfect calibration). Overall bias was analysed by computing the proportion of participants who were over-/under-efficacious or calibrated, for each performance outcome, and comparing the proportions of participants in these groups using Chi-square tests with a null hypothesis of equal distribution. Participants whose bias scores were within 10% of calibration were considered calibrated (Stankov & Lee, 2014).

Using calibration plots, with the subjective judgement plotted on one axis and the objective outcome plotted on the other, self-efficacy bias was compared against a hypothetical line indicating perfect calibration. Unlike any single calibration co-efficient, calibration plots enable us to gauge the prevalence of both over- and under-efficaciousness and provide an easily interpretable visual display of bias across performance levels (Hacker et al., 2008; Pieschl, 2009). For each performance task, within-subjects *t*-tests were conducted to determine the difference between self-efficacy and performance at each of the five performance outcome levels, following previous calibration research (e.g., Ehrlinger, Johnson, Banner, Dunning, & Kruger, 2008). To reduce the risk of familywise error, the alpha level for these *t*-tests was set at .01.

Two hierarchical regression analyses were used to explore how self-efficacy calibration bias and accuracy on the first report or exam related to performance on the second report or exam, respectively. As a known predictor of performance, task-based self-efficacy measured immediately prior to the time 2 performance outcome was entered at step 1, with self-efficacy bias and accuracy scores calculated from time 1 entered at step 2. Predictor variables were mean-centred. Cohort was also entered as a predictor variable to account for variability associated with cohort differences.

Results

Preliminary analyses

Dropout analyses were conducted using logistic regression with attrition versus completion of all study questionnaires predicted from the study variables of baseline self-efficacy and final grades. Participants completing all waves of data collection were more likely to have higher final grades (OR=1.04, p=.02).

A series of between-groups t-tests was conducted to compare the 2014/2015 cohort means on self-efficacy and performance variables. There was a significant difference between the two cohorts for grades on the second written report (2014: M=58.6, SD=14.1; 2015: M=64.2, SD=12.1). This was identified as an administrative error associated with marking procedures in 2015; therefore written report grades for the 2015 cohort were transformed to reflect the mean representing the institutional norm for that assessment. There were no other significant differences between the two cohorts.

Are students' self-efficacy beliefs calibrated or biased?

Mean self-efficacy and academic performance for the five performance outcomes are shown in Table 1, along with the correlation between self-efficacy and performance in each case. Small-to-medium significant positive correlations were observed. Table 1 also shows mean accuracy scores for the five performance outcomes. All were significantly different from zero (perfect calibration).

Table 1

Mean self-efficacy and grades, correlations between self-efficacy and performance, and mean accuracy of self-efficacy

	Report 1	Report 2	Exam 1	Exam 2	Subject grade
n	207	161	149	126	197
Self-efficacy	1.80 (.86)	1.84 (.85)	2.01 (.96)	1.88 (.90)	2.63 (.86)
Performance	1.82 (1.08)	1.76 (1.06)	2.96 (1.07)	1.96 (1.08)	1.62 (.92)
Correlation	.30***	.32***	.34***	.56***	.21**
Accuracy	.93 (.69)***	.91 (.66)***	1.30 (.81)***	.80 (.60)***	1.22 (.93)***

Note: Standard deviations are in parentheses; significance values for accuracy reflect one-sample *t*-tests of the difference between mean accuracy and zero (perfect calibration) ** $p<.01$ *** $p<.001$

Figure 2 shows the proportion of over-efficacious, calibrated, and under-efficacious participants for the five performance outcomes. For both written reports and the final exam, over- and under-efficaciousness and calibration were roughly equally prevalent, though there was a consistent slight tendency towards under-efficaciousness. For the mid-term exam, under-efficaciousness was pronounced, $\chi^2 = 121.30$, $p<.001$. Conversely, over-efficaciousness was prevalent in regard to overall subject performance, $\chi^2 = 136.46$, $p<.001$.

In the calibration plots (Figures 3–7) self-efficacy bias above the calibration line indicates over-efficaciousness, and bias below the line indicates under-efficaciousness. The plots indicate that, generally, weaker performers are over-efficacious while stronger performers are under-efficacious. The figures are annotated to show the results of the within-subjects *t*-tests comparing actual self-efficacy bias to calibration at each grade level ($\alpha=.01$).

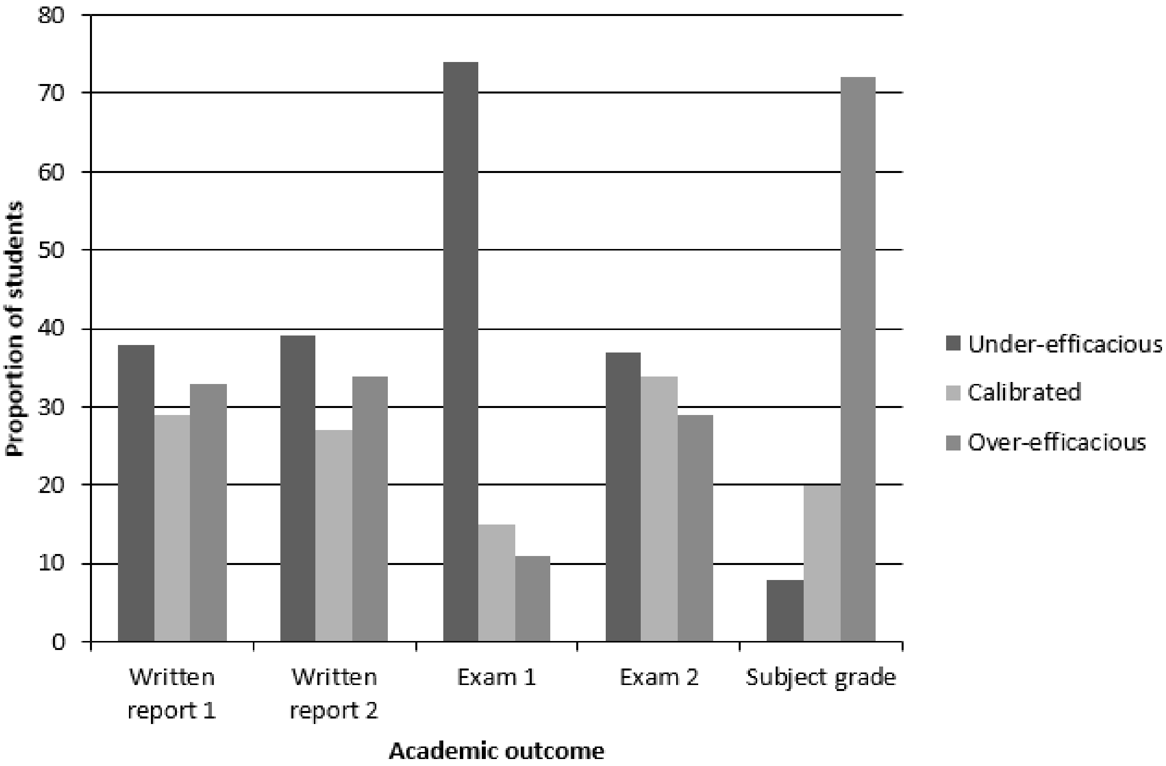


Figure 2. Self-efficacy calibration proportions by performance outcome

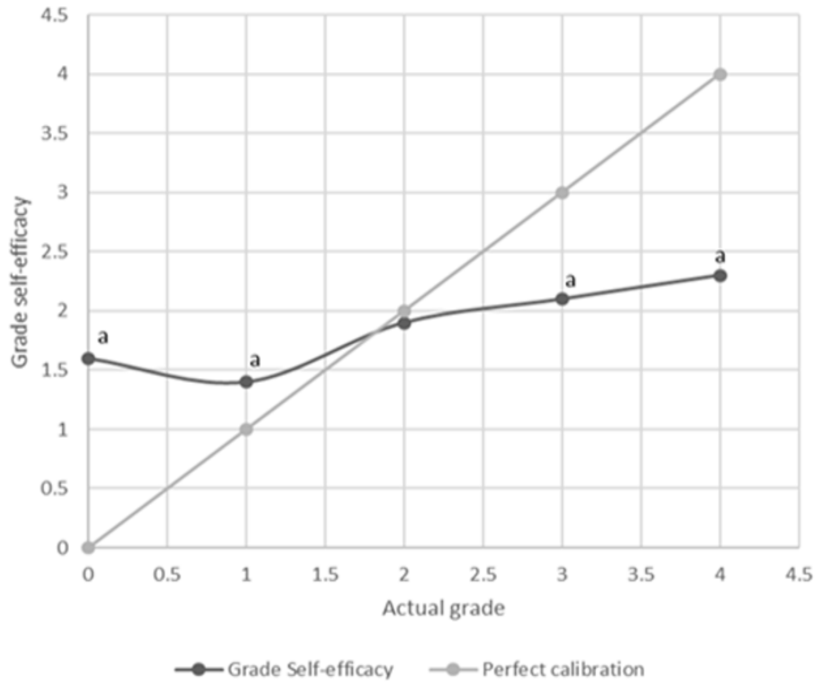


Figure 3. Self-efficacy calibration for written report 1

a = self-efficacy deviated from calibration at *p*<.001

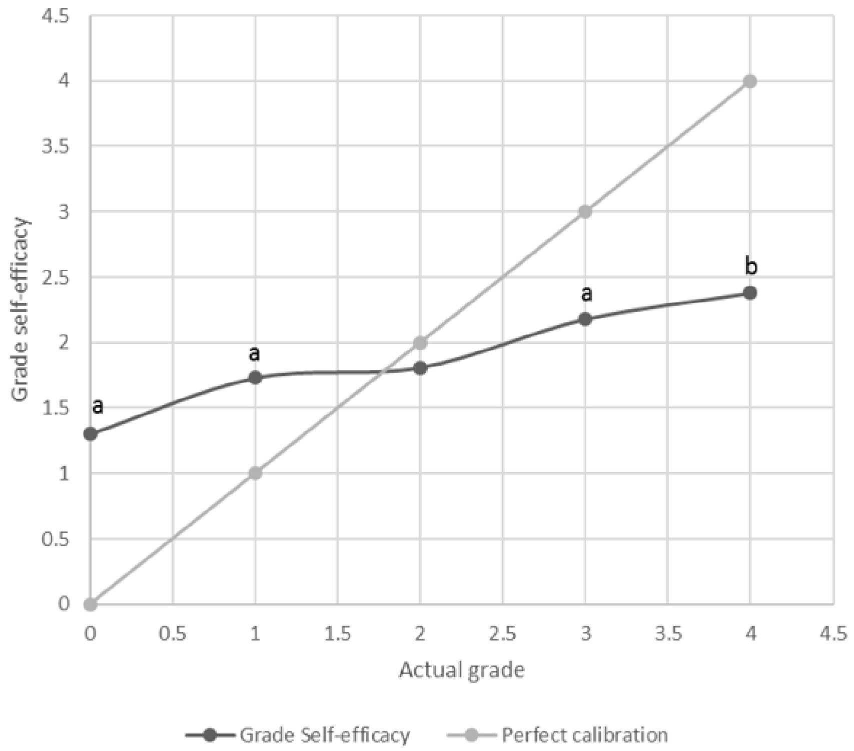


Figure 4. Self-efficacy calibration for written report 2

a = self-efficacy deviated from calibration at *p*<.001, b = self-efficacy deviated from calibration at *p*<.01

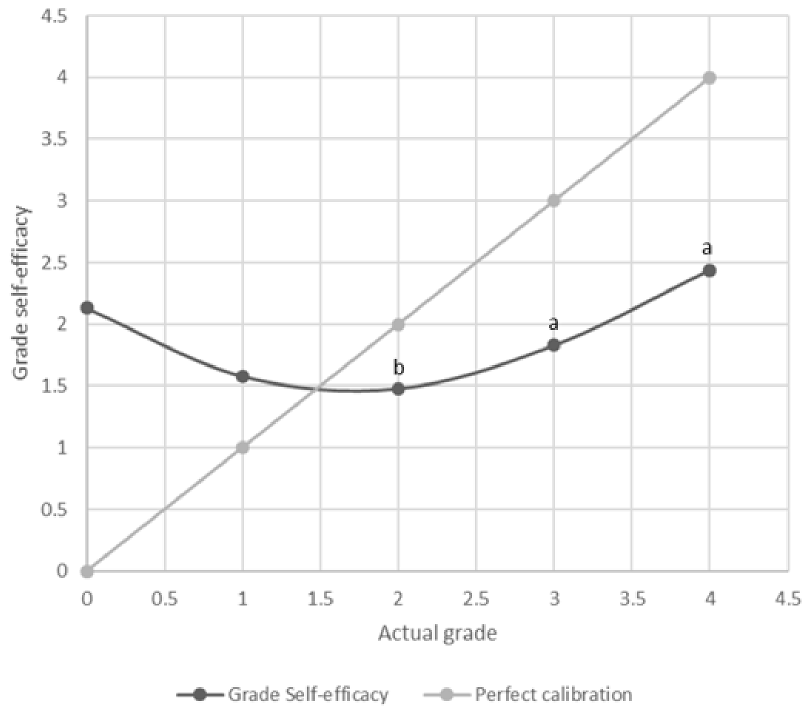


Figure 5. Self-efficacy calibration for mid-semester exam

a = self-efficacy deviated from calibration at $p<.001$, b = self-efficacy deviated from calibration at $p<.01$

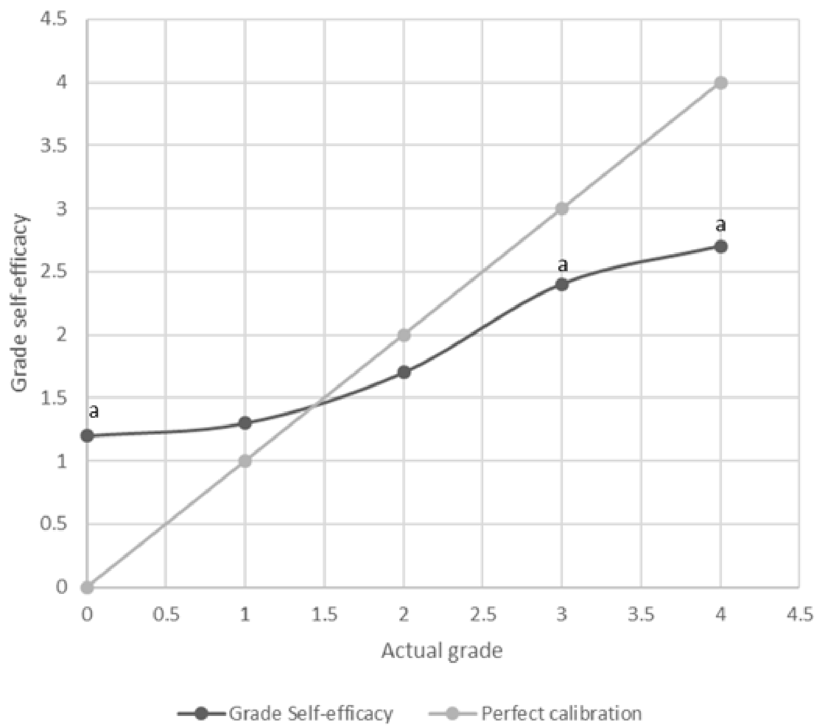


Figure 6. Self-efficacy calibration for end-of-semester exam

a = self-efficacy deviated from calibration at $p<.001$

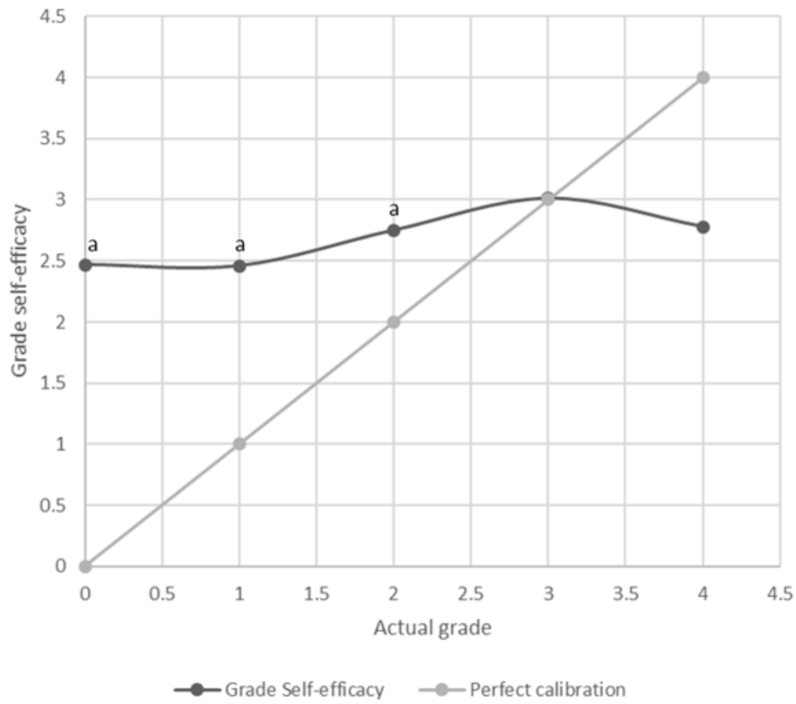


Figure 7. Self-efficacy calibration for overall subject grade

a = self-efficacy deviated from calibration at *p*<.001

What level of self-efficacy is optimal with regard to academic performance outcomes: over-efficaciousness, under-efficaciousness, or accuracy of self-efficacy beliefs?

Self-efficacy positively predicted performance, while self-efficacy bias negatively predicted performance (under-efficaciousness was associated with better performance), on both the written report and exam. Self-efficacy accuracy did not predict performance outcomes. Results of the hierarchical regression analyses are presented in Table 2.

Table 2

Hierarchical regression results for prediction of written report and exam performance

	Step	Variables entered	β	R Square	R Square change
Written report	1	Cohort	.046		
(n=163)		Self-efficacy	.288***	.084***	
	2	Cohort	.068		
		Self-efficacy	.290***		
		Bias	-.337***		
		Accuracy	-.111	.212***	.128***
Exam	1	Cohort	-.062		
(n=135)		Self-efficacy	.546***	.303***	
	2	Cohort	.091		
		Self-efficacy	.690***		
		Bias	-.402***		
		Accuracy	.049	.453***	.150***

Note: Self-efficacy was measured immediately prior to the performance outcome at time 2. Bias and accuracy were calculated based on measurements of self-efficacy and performance at time 1. *** $p \leq .001$

Discussion

This study examined the calibration (accuracy and bias) of tertiary students' self-efficacy beliefs with respect to actual performance outcomes. Overall, students' self-efficacy beliefs were poorly calibrated with actual capacity to perform. Self-efficacy beliefs were inaccurate overall for each of the performance outcomes. In terms of self-efficacy bias, under-efficaciousness was more prevalent at task level (written reports and exams) and over-efficaciousness was more prevalent for overall subject outcome. Weaker performers tended to be over-efficacious, while the reverse was true for stronger performers. Self-efficacy bias on one occasion predicted future performance on the same type of task, with under-efficacious students showing better performance capacity. Accuracy of self-efficacy beliefs did not predict performance outcomes.

Are students' self-efficacy beliefs calibrated or biased?

Analyses of self-efficacy accuracy showed that students' self-efficacy beliefs deviated significantly from their performance outcomes across all five measures. For individual academic tasks (written reports and exams), under-efficaciousness was at least as prevalent as over-efficaciousness, accounting for more than a third of students for each task. This is in contrast with previous research that found the majority of students to be over-efficacious (Chen & Zimmerman, 2007; Chen, 2003; Ramdass & Zimmerman, 2008). However, a different trend applied to a broader-level judgement that was more temporally distal: more than two-thirds of participants' self-efficacy beliefs for overall subject grades exceeded their objective performance capacity, and less than ten percent were under-efficacious.

Calibration plots and tests of differences between self-efficacy and performance at different performance levels illustrated a consistent trend for weaker performers to be over-efficacious, and for stronger performers to be under-efficacious. While the magnitude of this trend varied across tasks, it was evident across all performance outcomes. While this pattern is consistent with previous research (Bol et al., 2005; Hacker et al., 2008), our findings suggest equal miscalibration for over- and underperformers whereas the previous studies suggest stronger miscalibration in underperformers.

What level of self-efficacy is optimal with regard to academic performance outcomes: over-efficaciousness, under-efficaciousness, or accuracy of self-efficacy beliefs?

On all five measures, self-efficacy was significantly positively correlated with performance outcomes, with small-to-medium effects. This is consistent with previous findings in the literature. However, correlational analyses can only tell part of the story. While it is true that the strongest students' self-efficacy beliefs consistently exceeded those of the poorest students (reflected in the positive correlations and visible in the calibration plots), the self-efficacy beliefs of the strongest students also consistently fell short of their own capacity to perform, while weaker students by comparison were over-efficacious. The key issue is the point of comparison: low-achievers' self-efficacy beliefs are somewhat low when compared to high-achievers – but, they are actually high when compared to their own performance outcomes. Thus, the positive correlation between self-efficacy and academic performance appears to obscure consistent trends in self-efficacy bias.

The regression results suggest that under-efficaciousness is most adaptive, with self-efficacy bias tending towards under-efficaciousness being associated with stronger performance on subsequent analogous written report and exam tasks. The effects noted were medium-to-large by convention. Self-efficacy accuracy did not significantly predict subsequent performance outcomes on either the written report or exam. This contrasts with findings that greater accuracy is related to better performance in mathematics for school-aged students (Chen, 2003). As we included both bias and accuracy in our analyses, we were able to pick up the distinction between over- and under-efficaciousness, both of which would have been reflected as non-directional inaccuracy in previous studies.

On face value, these findings are inconsistent with recommendations growing out of self-efficacy theory (Bandura, 1997; Pajares, 2006) that self-efficacy beliefs which exceed capacity are predictive of better outcomes. In contrast, they suggest that the motivation to reduce the discrepancy between the perceived current state of ability and the desired state of ability leads to increased effort and improved performance (e.g., Carver & Scheier, 1982; Vancouver & Kendall, 2006). However, even those who suggest that such an interpretation is plausible caution against translating this into practical interventions designed to increase under-efficaciousness, citing the potential risk of disengagement (Vancouver & Kendall, 2006).

Critically, the above interpretation rests on an assumption about causality. The measurement of bias prior to the measurement of performance provides some indication that it is not performance that directly influences bias (and we would not logically expect strong performance to directly result in under-efficaciousness). However, it seems perhaps most plausible that both self-efficacy bias and performance are related to a cocktail of other exogenous variables. We turn to a discussion of these variables here, focusing on three main trends in turn, namely, over-efficaciousness in weaker performers, under-efficaciousness in stronger performers, and strong over-efficaciousness at subject grade level.

Over-efficaciousness in weaker performers

A number of potential explanations for meta-cognitive overconfidence have been proffered, many of which are also plausible explanations for over-efficaciousness in weaker students. One example is the "unskilled and unaware" explanation, which suggests that poorer performers suffer a dual burden – the lack of capacity to perform well also manifests as a lack of insight into what is required to perform well, resulting in over-confidence (Kruger & Dunning, 1999). Similarly, weaker students' sense of efficacy may exceed their capacity to perform because they are unaware of where they are lacking – in this case, "unable and unaware". Other explanations based on motivated biases include the better-than-average effect (Krueger & Mueller, 2002), self-enhancement biases (Ehrlinger, Mitchum, & Dweck, 2016) and defensive self-deception (Stankov & Lee 2014).

Under-efficaciousness in stronger performers

Most previous research has discussed overconfidence, thus only little research attention has been paid to the potential underpinnings of under-confidence. However, several explanations are possible. For example, the self-efficacy beliefs of strong performers may be subject to defensive pessimism (Norem & Cantor, 1986) or "bracing for the worst", whereby people endeavour to protect themselves from the disappointment of negative results or feedback by lowering their outcome expectations (e.g., Taylor & Shepperd, 1998). This tendency is observed for self-relevant outcomes according to how detrimental an anticipated loss is perceived to be (Shepperd, Findley-Klein, Kwavnick, Walker, & Perez, 2000); thus, stronger students may be more likely to brace for loss given the greater likelihood that their self-concepts and self-esteem rest on strong academic performance (Higgins, 1987).

Recent qualitative research also suggests that individuals with stronger cognitive abilities may simply be more humble or cautious in their self-evaluations, preferring to provide lower judgements of confidence in their ability to perform (de Carvalho Filho, 2009). Students who underestimate their performance capacity report not wanting to "jinx" themselves and preferring to "play it safe" (Bol et al., 2005, p. 272). Strong students may also wish, for social reasons, to avoid appearing too competent or hard-working (Schunk & Pajares, 2004).

In observing over-efficaciousness in weaker performers and under-efficaciousness in stronger performers, in terms of task-based self-efficacy, it is also interesting to note that those who performed at an average level (i.e., those who achieved around a strong pass or credit grade) appeared to hold the most accurate self-efficacy beliefs. One might speculate that average students are less subject to the types of biases which influence the self-efficacy beliefs of those students at the more extreme ends of the academic performance continuum.

Over-efficaciousness for subject grades

While task-level self-efficacy calibration was not directly compared to subject-level in the present case, the statistics reported above converge on the conclusion that the grossest deviation from perfect calibration occurred at subject-level. This is consistent with other studies' findings of higher self-estimates at a broader level compared with a more specific level (Ackerman et al., 2002). The broader the domain under investigation, the more abstract self-judgements become, with the effect that they are more prone to influence from self-serving criteria (Dunning, Johnson, Ehrlinger, & Kruger, 2003).

A critical issue in terms of over-efficaciousness regarding overall subject grades relates to the timing of measurement. In the present study, self-efficacy for subject grades was measured close to the beginning of semester; thus, subject-level self-efficacy judgements were made largely in the absence of relevant mastery information, a key source of self-efficacy beliefs (Bandura, 1997). A further timing issue is that the influences which may reduce overconfidence, such as defensive pessimism, are more likely to be at play as the potentially sobering outcome or feedback approaches, whereas temporally distant outcomes – such as overall grades – are more likely to be subject to unrealistic optimism (Sweeny & Krizan, 2013; Taylor & Shepperd, 1998).

When students make self-efficacy judgements for broader-level outcomes that are distant in time, they may use inappropriate anchors for their self-efficacy beliefs and insufficiently adjust from that point (Bol & Hacker, 2012; Bol et al., 2005). For example, self-efficacy may be based on desired outcomes or aspirations (Serra & DeMarree, 2016) or norms such as average GPA of an institution (Clayson, 2005).

Limitations

The drop-out analysis identified that participants who completed all of the questionnaires were more likely to be better performers. This provides some evidence of systematic attrition in the present study, with poorer performers under-represented. In addition, as noted in the results section, there were differences in performance on the written report between the two cohorts included. Steps were taken to control for cohort differences.

A further limitation relates to the interpretation of substantial under-efficaciousness for the first exam. Inspection of the mean performance and self-efficacy ratings suggests that participants may appear so under-efficacious for this task because it was easier than anticipated. That is, under-efficaciousness in this case appears to reflect the fact that students performed comparatively well on this task, not that they had particularly weak self-efficacy beliefs.

Directions for future research

A strength of the present study was the analysis of self-efficacy calibration across different task types and broader levels of outcome specificity not previously explored in self-efficacy calibration research. The findings highlight a potential lack of consistency in self-efficacy calibration depending on the performance outcome being investigated, and are suggestive of an interaction effect between ability level and performance task. Future research directly comparing calibration within-subjects across tasks and domains is warranted. Previous research suggests that interventions may improve the accuracy of mathematics self-efficacy beliefs in school-aged children (Ramdass & Zimmerman, 2008); further research into whether such interventions improve the accuracy of university students' self-efficacy beliefs is warranted. Furthermore, little is known about what differentiates over-efficacious and under-efficacious students (Ehrlinger et al., 2016). Identifying characteristics of biased students is potentially a fruitful avenue for further applied research.

In regard to the use of correlation coefficients and absolute accuracy as measures of calibration, it is evident from Table 1 and the calibration plots that very similar correlation coefficients can show a very divergent distribution of over- and under-efficacious individuals across ability levels. In the same vein, focusing only on accuracy would have obscured some important observations in the present case. We recommend that researchers exploring calibration carefully consider the range of measures available (Dunning & Helzer, 2014).

Conclusion

The findings of this study suggest that many university students' self-efficacy beliefs for academic performance are miscalibrated, with inaccuracy manifesting in both under- and over-efficaciousness. In a result that may surprise educators, educational researchers and policy makers, it was under-efficaciousness, rather than over-efficaciousness or accuracy of self-efficacy beliefs, that predicted future performance outcomes in this sample. We replicated the commonly-reported finding that higher levels of self-efficacy were generally related to stronger performance outcomes; however, this finding cannot be interpreted in isolation. A key point is that, although low-achievers' self-efficacy beliefs did fall somewhat short of those of their high-achieving peers, they also were consistently higher than their own performance outcomes would warrant – that is, low-achievers were over-efficacious across all performance outcomes. Conversely, under-efficaciousness was prevalent in stronger performers. These findings cast some doubt on the notion of self-efficacy as self-fulfilling prophecy. In the simplest terms, we see that many students who believe they can, actually cannot, and many students who believe they cannot, actually can. Rather than identifying under-efficaciousness as a risk factor for poor performance, our data support Shakespeare's assertion that modest doubt is a beacon of the wise.

Blanket recommendations to enhance self-efficacy continue to proliferate (e.g., Moores et al., 2006). And yet, according to our results, it is unclear who would benefit from this approach, given that those under-efficacious students most in need of having their self-efficacy beliefs bolstered are those who are least in need of performance improvements. Meanwhile, the poorest performers appear to be in little danger of making self-debilitating judgements of their performance capacity. Further boosting the self-efficacy beliefs of low-achieving tertiary students appears unlikely to be of benefit when these students appear to suffer no lack of self-efficacy when considering their actual performance capabilities. Instead, the over-efficaciousness observed in weaker performers puts them at risk of those potential negative impacts discussed earlier (Boekaerts & Rozendaal, 2010; Vancouver & Kendall, 2006), such as stopping studying before they are properly prepared for an assessment, or refraining from seeking academic support which is sorely needed. The provision of accurate formative feedback, including negative feedback and constructive criticism where applicable, may promote more realistic self-efficacy beliefs.

References

Ackerman, P. L., Beier, M. E., & Bowen, K. R. (2002). What we really know about our abilities and our knowledge. *Personality and Individual Differences, 33*(4), 587-605. doi:10.1016/S0191-8869(01)00174-X

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 5, pp. 307-337). Greenwich, CT: Information Age.

Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology, 88*(1), 87-99. doi:10.1037/0021-9010.88.1.87

Boekaerts, M., & Rozendaal, J. S. (2010). Using multiple calibration indices in order to capture the complex picture of what affects students' accuracy of feeling of confidence. *Learning and Instruction, 20*(5), 372-382. doi:10.1016/j.learninstruc.2009.03.002

Bol, L., & Hacker, D. J. (2012). Calibration research: Where do we go from here? *Frontiers in Psychology, 3*, 229. doi:10.3389/fpsyg.2012.00229

Bol, L., Hacker, D. J., O'Shea, P., & Allen, D. (2005). The influence of overt practice, achievement level, and explanatory style on calibration accuracy and performance. *The Journal of Experimental Education, 73*(4), 269-290. doi:10.3200/JEXE.73.4.269-290

Carver, C. S., & Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality-social, clinical, and health psychology. *Psychological Bulletin, 92*(1), 111-135. doi:10.1037/0033-2909.92.1.111

Chen, P., & Zimmerman, B. (2007). A cross-national comparison study on the accuracy of self-efficacy beliefs of middle-school mathematics students. *The Journal of Experimental Education, 75*(3), 221-244. doi:10.3200/JEXE.75.3.221-244

Chen, P. P. (2003). Exploring the accuracy and predictability of the self-efficacy beliefs of seventh-grade mathematics students. *Learning and Individual Differences, 14*(1), 77-90. doi:10.1016/j.lindif.2003.08.003

Clayson, D. E. (2005). Performance overconfidence: Metacognitive effects or misplaced student expectations? *Journal of Marketing Education, 27*(2), 122-129. doi:10.1177/0273475304273525

de Carvalho Filho, M. K. (2009). Confidence judgments in real classroom settings: Monitoring performance in different types of tests. *International journal of psychology, 44*(2), 93-108. doi:10.1080/00207590701436744

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction, 22*(4), 271-280. doi:10.1016/j.learninstruc.2011.08.003

Dunlosky, J., & Thiede, K. W. (2013). Four cornerstones of calibration research: Why understanding students' judgments can improve their achievement. *Learning and Instruction, 24*, 58-61. doi:10.1016/j.learninstruc.2012.05.002

Dunning, D., & Helzer, E. G. (2014). Beyond the correlation coefficient in studies of self-assessment accuracy: Commentary on zell & krizan (2014). *Perspectives on Psychological Science, 9*(2), 126-130. doi:10.1177/1745691614521244

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science, 12*(3), 83-87. doi:10.1111/1467-8721.01235

Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes, 105*(1), 98-121. doi:10.1016/j.obhdp.2007.05.002

Ehrlinger, J., Mitchum, A. L., & Dweck, C. S. (2016). Understanding overconfidence: Theories of intelligence, preferential attention, and distorted self-assessment. *Journal of Experimental Social Psychology*, *63*, 94-100. doi:10.1016/j.jesp.2015.11.001

Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management Review, 17*(2), 183-211. doi:10.5465/AMR.1992.4279530

Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology, 92*(1), 160-170. doi:10.1037/0022-0663.92.1.160

Hacker, D. J., Bol, L., & Keener, M. C. (2008). Metacognition in education: A focus on calibration. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 429-455). New York: Psychology Press.

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94*(3), 319-340. doi:10.1037/0033-295X.94.3.319

Krueger, J., & Mueller, R. A. (2002). Unskilled, unaware, or both? The better-than-average heuristic and statistical regression predict errors in estimates of own performance. *Journal of Personality and Social Psychology*, *82*(2), 180-188. doi:10.1037/0022-3514.82.2.180

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of*

Personality and Social Psychology, 77(6), 1121-1134. doi:10.1037/0022-3514.77.6.1121

Lee, C., & Bobko, P. (1994). Self-efficacy beliefs: Comparison of five measures. *Journal of Applied Psychology, 79*(3), 364. doi:10.1037/0021-9010.79.3.364

Mengelkamp, C., & Bannert, M. (2009). Judgements about knowledge: Searching for factors that influence their validity. *Electronic Journal of Research in Educational Psychology, 7*(17), 163-190.

Moores, T. T., Chang, J. C.-J., & Smith, D. K. (2006). Clarifying the role of self-efficacy and metacognition as predictors of performance: Construct development and test. *ACM SIGMIS Database, 37*(2-3), 125-132. doi:10.1145/1161345.1161360

Norem, J. K., & Cantor, N. (1986). Anticipatory and post hoc cushioning strategies: Optimism and defensive pessimism in “risky” situations. *Cognitive Therapy and Research, 10*(3), 347-362. doi:10.1007/bf01173471

Pajares, F. (2006). Self-efficacy during childhood and adolescence: Implications for teachers and parents. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 339-368). Greenwich, CT: Information Age.

Pajares, F., & Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. *Contemporary Educational Psychology, 24*(2), 124-139. doi:10.1006/ceps.1998.0991

Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. *Contemporary Educational Psychology, 20*(4), 426-443. doi:10.1006/ceps.1995.1029

Pieschl, S. (2009). Metacognitive calibration—an extended conceptualization and potential applications. *Metacognition and Learning, 4*(1), 3-31. doi:10.1007/s11409-008-9030-4

Ramdass, D., & Zimmerman, B. J. (2008). Effects of self-correction strategy training on middle school students' self-efficacy, self-evaluation, and mathematics division learning. *Journal of Advanced Academics, 20*(1), 18-41. doi:10.4219/jaa-2008-869

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353-387. doi:10.1037/a0026838

Schraw, G. (2009). Measuring metacognitive judgments. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 415-429). New York: Routledge.

Schunk, D. H., & Pajares, F. (2004). Self-efficacy in education revisited. In D. M. McInerney & S. Van Etten (Eds.), *Big theories revisited* (Vol. 4, pp. 115-138). Greenwich, CT: Information Age Publishing.

Serra, M. J., & DeMarree, K. G. (2016). Unskilled and unaware in the classroom: College students' desired grades predict their biased grade predictions. *Memory & Cognition, 44*(7), 1127-1137. doi:10.3758/s13421-016-0624-9

Shepperd, J. A., Findley-Klein, C., Kwavnick, K. D., Walker, D., & Perez, S. (2000). Bracing for loss. *Journal of Personality and Social Psychology, 78*(4), 620-634.

Stankov, L., & Lee, J. (2014). Overconfidence across world regions. *Journal of Cross-Cultural Psychology, 45*(5), 821-837. doi:10.1177/0022022114527345

Stankov, L., & Lee, J. (2017). Self-beliefs: Strong correlates of mathematics achievement and intelligence. *Intelligence, 61*, 11-16. doi:10.1016/j.intell.2016.12.001

Sweeny, K., & Krizan, Z. (2013). Sobering up: A quantitative review of temporal declines in expectations. *Psychological Bulletin, 139*(3), 702-724. doi:10.1037/a0029951

Taylor, K. M., & Shepperd, J. A. (1998). Bracing for the worst: Severity, testing, and feedback timing as moderators of the optimistic bias. *Personality and Social Psychology Bulletin, 24*(9), 915-926. doi:doi:10.1177/0146167298249001

Usher, E. (2016). Personal capability beliefs. In L. Corno & E. M. Anderman (Eds.), *Handbook of educational psychology* (Vol. 3, pp. 146-159). New York: Routledge.

Vancouver, J. B., & Kendall, L. N. (2006). When self-efficacy negatively relates to motivation and performance in a learning context. *Journal of Applied Psychology, 91*(5), 1146-1153. doi:10.1037/0021-9010.91.5.1146

Zabrucky, K. M. (2010). Knowing what we know and do not know: Educational and real world implications. *Procedia: Social and Behavioral Sciences*, *2*(2), 1266-1269. doi:10.1016/j.sbspro.2010.03.185

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into practice, 41*(2), 64-70. doi:10.1207/s15430421tip4102_2

Zvacek, S. M., de Fátima Chouzal, M., & Restivo, M. T. (2015). Accuracy of self-assessment among graduate students in mechanical engineering. *Proceedings of the International Conference on Interactive Collaborative Learning (ICL), 2015*, 1130-1133. doi:10.1109/ICL.2015.7318192

Chapter 5: Study 4

When belief exceeds capacity:

Which factors differentiate over-efficacious students from their peers?

Abstract

Background: Research indicates that overconfident students, such as those with positive self-efficacy bias (i.e., over-efficaciousness), are at risk of poor academic outcomes. However, little is known about what characterises over-efficacious students. In this study, we examine whether individual differences (age, sex, cognitive ability, self-beliefs, personality traits) and contextual factors (SES, educational background) differentiate over-efficacious students from their peers.

Method: In a longitudinal design, first-year university students' (n=197) self-efficacy beliefs were assessed at baseline (T1), mid-semester (T2), and end-of-semester (before exams; T3). Self-efficacy scores were compared to final subject grades (T4). Predictor variable data were collected at T1. Participants were identified as over-efficacious (self-efficacy exceeded performance by half a grade band or more) or non over-efficacious at each time point.

Results: Logistic regression analyses indicated that, compared to their peers, over-efficacious students were: lower in cognitive ability (OR=.80) and agreeableness (OR=.88) (T1); younger (OR=.96) with higher self-esteem (OR=1.13) (T2); and, higher in self-esteem (OR=1.11) and SES (1.01) (T3).

Conclusions: Key results suggest an initial influence of cognitive processes and a sustained influence of motivated self-evaluation processes on illusory self-efficacy beliefs. Findings may inform educational policies for pre-tertiary students and interventions for university students likely to be over-efficacious, reducing associated academic risks.

Key words: *self-efficacy, academic performance, calibration, bias, confidence*

A lack of correspondence between a person's actual ability and their confidence in that ability is a robust finding across many domains (Dunning, 2005). In educational settings, calibration research similarly shows that learners' ideas about themselves and their abilities are often disconnected from reality. Whether confidence is measured as a prediction of academic performance (Ehrlinger, Johnson, Banner, Dunning, & Kruger, 2008), as academic self-concept (Sheldrake, 2016), or academic self-efficacy (Chen, 2003; Talsma, Norris, & Schuz, 2017), research consistently shows substantial miscalibration between learners' subjective judgements and actual academic outcomes.

A key concern is when belief exceeds capacity, which is the most common finding in calibration research (Hacker, Bol, & Keener, 2008). Overconfidence is argued to undermine academic self-regulation, potentially leading to premature termination of study, complacency, decreased help-seeking, and ultimately, poor performance outcomes (e.g., Boekaerts & Rozendaal, 2010; Dunlosky & Rawson, 2012; Talsma, Norris, et al., 2017). Thus, knowing what characterises individuals whose subjective judgements exceed their objective capacity could inform educational policies designed to improve academic performance (Sheldrake, 2016). However, little is known about what predicts confidence bias – that is, under- or overconfidence (Ehrlinger, Mitchum, & Dweck, 2016; Sheldrake, 2016). The present study aims to address this gap in the literature, focusing on academic self-efficacy bias, to investigate what differentiates over-efficacious students from their peers.

Self-efficacy – perceptions of one's capability to organise and execute required courses of action to achieve particular outcomes (Bandura, 1997) – is widely believed to be one of the most important non-intellectual determinants of academic performance (Richardson, Abraham, & Bond, 2012). While self-efficacy has been shown to be positively related to academic achievement (Talsma, Schüz, Schwarzer, & Norris, 2017), research also indicates that self-efficacy beliefs do not always align with capacity to perform (Chen, 2003;

Talsma, Norris, et al., 2017). Self-efficacy bias, the magnitude and direction of the deviation between self-efficacy and academic performance, may be positive (over-efficaciousness) or negative (under-efficaciousness)(Hacker et al., 2008).

A comprehensive theory regarding variables which would differentially predict over-efficaciousness has not yet been developed. Self-efficacy is situated within the triadic reciprocal determinism framework of social cognitive theory, which would suggest that both person factors and environmental factors may be implicated (Bandura, 1997). In this exploratory context, we approach the issue by conducting logistic regression analyses using backwards elimination, and we provide the following rationale for the inclusion of predictor variables.

A review of the literature suggests two important themes warrant attention when considering what might distinguish over-efficacious individuals from their peers: cognitive biases and motivated biases (Stankov & Lee, 2014). Cognitive biases might account for inaccurate calibration because of faulty judgement processes associated with the dual burden argued to be suffered by those who are simultaneously unskilled and unaware of it (Ehrlinger et al., 2008; Kruger & Dunning, 1999). In this view, individuals who lack competence in a given domain inevitably lack insight into what is required to perform successfully, which leads to over-confidence. Thus, we propose that participants with lower scores on a test of cognitive ability will be more likely to be over-efficacious.

In terms of motivated biases, it has long been suggested that psychological benefits can be obtained from positive self-views (Taylor & Brown, 1988). Overconfidence has been described as a motivated drive towards self-enhancement (Ehrlinger et al., 2016), with individuals with higher self-esteem more likely to adopt illusorily positive self views in a range of areas (Baumeister, Campbell, Krueger, & Vohs, 2003). Owing to the positive self-image enjoyed by those with high self-esteem, further “esteem-enhancing illusions” which would serve to maintain high self-esteem or protect self-concept are unreservedly adopted (Brown, Collins, & Schmidt, 1988, p. 446). For this reason, those with positive global-level self-views such as self-esteem, generalised self-efficacy, and academic self-concept may be more likely to be over-efficacious.

Five Factor Model (FFM) personality traits may also be implicated in over-efficaciousness. Example items for the FFM traits of openness to experience, conscientiousness, extraversion, agreeableness and neuroticism (McCrae & Costa, 1987) are presented in Table 1. FFM traits have been shown to relate to the realism of confidence judgements for general knowledge questions and cognitive tests (Dahl, Allwood, Rennemark, & Hagberg, 2010; Pallier et al., 2002). While FFM traits are associated both with self-efficacy and with academic outcomes (Caprara, Vecchione, Alessandri, Gerbino, & Barbaranelli, 2011; De Feyter, Caers, Vigna, & Berings, 2012), research regarding the influence of personality on self-efficacy bias specifically is limited. In this study, we explore whether FFM traits distinguish over-efficacious students from their peers.

There are mixed findings in the literature regarding the relationship between age and the accuracy of subjective judgements (e.g., de Bruin, Parker, & Fischhoff, 2012). However, older students appear to have more accurate perceptions of their abilities than younger students (e.g., Grimes, 2002). Thus, we suggest that younger students are more likely to demonstrate over-efficaciousness. With regard to participant sex, evidence across a number of domains suggests that males are more likely to have inaccurately positive self-evaluations (e.g., Sheldrake, Mujtaba, & Reiss, 2014); therefore, we predict that males will be more likely than females to be over-efficacious.

Turning to potential environmental influences on self-efficacy bias, research suggests that first-generation tertiary students are at risk of more negative self-views than their peers (Carpenter & Clayton, 2014) but do not necessarily perform more poorly (Inman & Mayes, 1999); thus, they may be more likely to be under-efficacious than their peers. Students from high SES families perform more strongly academically and are more confident, while lower SES students report more doubts about their academic abilities (Bloom, 2007; Neuenschwander, Vida, Garrett, & Eccles, 2007; Sirin, 2005). These findings suggest that both higher and lower SES students may be accurate in their beliefs because of a relative match between subjective beliefs and objective outcomes. However, as little research addresses SES and confidence bias specifically, we refrain from making a directional hypothesis.

A better understanding of the personal and contextual characteristics which distinguish over-efficacious students from their peers may have implications for theory development as well as utility in the development of interventions aimed at reducing potentially harmful over-efficaciousness in learning settings. In accordance with Figure 1, we consider three waves of self-efficacy data with a single end-of-semester performance outcome, in order to explore whether different variables predict over-efficaciousness at different points in a course of study. Evidence of what factors are at play at different times of semester may help to inform targeted interventions which can be tailored to those most in need, in the hope of improving calibration of self-efficacy and academic performance, and preventing the potential risks associated with over-efficaciousness.

Method

Participants and procedure

First-year undergraduate psychology students at an Australian university participated in the study (*n*=197, 144 females, mean age 24.8 years, age range 17-66). Participants received course credit for participation. Ethics approval was obtained from the Tasmanian Social Sciences Human Research Ethics Committee. Data were collected over a 13-week semester via online questionnaires and from institutional records post-semester, in two cohorts, 2014 (n=95) and 2015 (n=102). Figure 1 illustrates the study design. Participants received study information and provided consent to participate through the online questionnaire instrument at T1. Mid-semester data collection (T2) occurred prior to the submission of any assessments or the receipt of any formalised feedback, while the end-of-semester questionnaire (T3) was completed subsequent to assignment feedback during the semester, but prior to completion of the final exam and receipt of final results (T4).

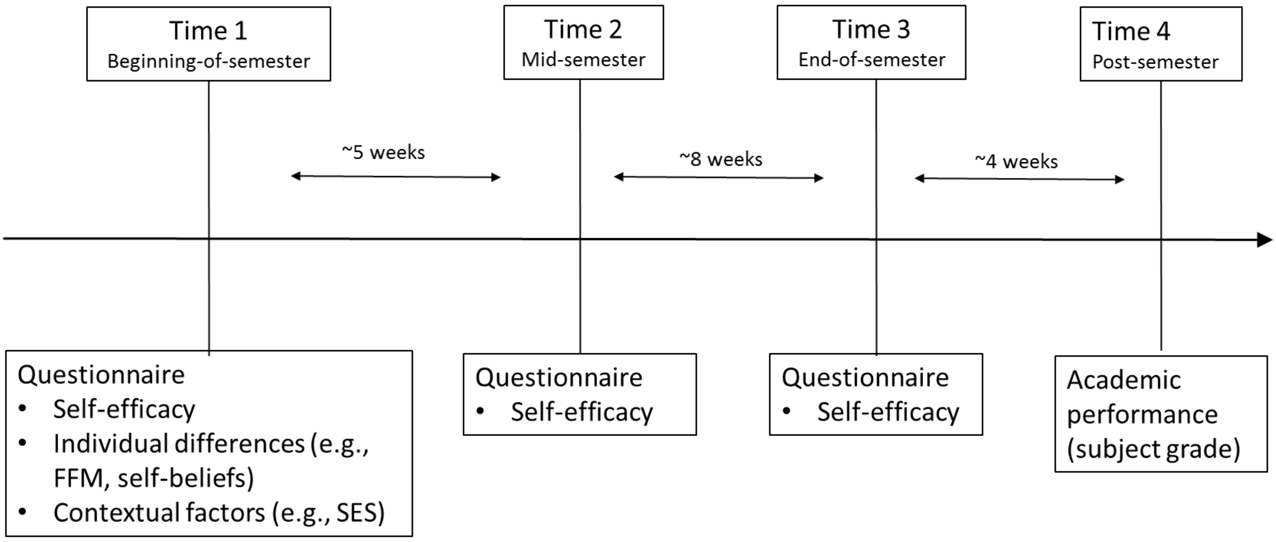


Figure 1. Study design

Note. Time between waves is approximate because questionnaires were completed online within a 5-day window.

Measures

Academic performance data was collected at T4. Academic performance was measured following Australian standards as the grade band achieved for the subject overall, with a possible range of 0 to 4 (fail [<50%], pass [50-59%], credit [60-69%], distinction [70-79%], and high distinction [≥80%], respectively). To ensure reliability, institutional policy provides that individual assessments from which the subject grade is formed are marked using assessment rubrics and moderated for consistency. *Academic self-efficacy (ASE)* was measured at T1-3 by self-report, using a composite measure of self efficacy magnitude and strength measured on the same scale as academic performance. Self-efficacy magnitude reflected participants' binary (yes=1/no=0) belief in their ability to achieve each of the academic performance grade bands, presented in order of increasing difficulty. Self-efficacy strength reflected participants' level of confidence (0-100%) in their ability to achieve each grade band, again with performance outcomes presented in increasing order of difficulty. Strength scores, as proportions, were summed for each "yes" decision at the magnitude level,. Thus, a person who responded "yes" they could achieve a pass (magnitude), and were 100% confident of this (strength) and then responded "yes" they could achieve a credit (magnitude) and were 70% confident of this (strength) but then responded "no" they could not achieve a distinction would have ASE score of 1.7. These procedures follow guidelines for self-efficacy measurement (Bandura, 2006; Lee & Bobko, 1994). *Cognitive ability* was measured using a time-limited 10-item logical reasoning test where participants were required to draw inferences from provided information (Ekstrom, French, Harman, & Dermen, 1976). Scale scores ranged from 0 to 10. Instrument authors report that internal consistency of the measure is α=.76-.78. *Educational background* was a categorical variable differentiating between first-in-family tertiary students and those with parents/grandparents who had attended university. An "I don't know" option was provided. *SES* was measured by applying the 2011 Australian Bureau of Statistics' socio-economic indexes for areas rankings using participants' main residential postcode. Participants' *sex, age* and *cohort* were recorded.

The following standardised survey instruments were used (See Table 1 for details). *Generalised self-efficacy (GSE)* was measured with the Generalised Self-efficacy Scale (Schwarzer & Jerusalem, 1995). *Academic self-concept (ASC)* was measured using the Academic Self-Concept Scale (Marsh, 1990). *Self-esteem* was measured using the Rosenberg scale (Rosenberg, 1965). *Domain self-efficacy (DSE)* was measured using the Motivated Strategies for Learning Questionnaire (Pintrich, Smith, Garcia, & McKeachie, 1991). FFM traits of *extraversion*, *agreeableness*, *neuroticism*, *conscientiousness* and *openness to experience* were measured using the Mini International Personality Item Pool scale (Donnellan, Oswald, Baird, & Lucas, 2006).

Table 1

Details of standardised instruments used

	No. of items	Example item	Scale points	Score descriptors	Score range	α
GSE	10	I can usually handle whatever comes my way	4	Not at all true: Exactly true	10-40	.91
ASC	6	Compared to my peers, I am good at psychology	6	False: True	6-36	.88
Self-esteem	10	I am able to do things as well as most other people	4	Strongly disagree: Strongly Agree	10-40	.89
DSE	8	I'm certain I can master the skills being taught in this class	7	Not at all true of me: Very true of me	8-56	.96
Extraversion	4	I am the life of the party	5	Very inaccurate: Very accurate	5-20	.81
Agreeableness	4	I sympathise with others' feelings	5	Very inaccurate: Very accurate	5-20	.76
Neuroticism	4	I have frequent mood swings	5	Very inaccurate: Very accurate	5-20	.69
Conscientiousness	4	I get chores done right away	5	Very inaccurate: Very accurate	5-20	.65
Openness to experience	4	I have a vivid imagination	5	Very inaccurate: Very accurate	5-20	.73

Note. GSE=generalised self-efficacy; ASC=academic self-concept; DSE=domain self-efficacy, α = Crobach's alpha

Analyses

In order to classify participants as over-efficacious or otherwise, we first calculated a self-efficacy bias score for each time point, as the signed deviation of the ASE score from the academic performance score. Positive values indicated over-efficaciousness, and vice versa. Participants were identified as over-efficacious if ASE exceeded academic performance by ≥0.5 (half a grade band), a cut-off which acknowledges that the range of scores which represent relative accuracy of self-beliefs may include a small amount of overconfidence (Sheldrake et al., 2014; Stankov & Lee, 2014). All other participants were classified as non over-efficacious. We took the present approach in order to apply a practically meaningful rather than statistical categorisation of participants into groups.

Logistic regression analyses were conducted to predict group membership (1: over-efficacious, 0: non over-efficacious) using backward elimination with the removal criterion based on likelihood ratios set at 0.05. Analyses controlled for cohort and DSE. Including DSE in the model enabled the exploration of variables which differentially predict over-efficaciousness and non-over-efficaciousness with self-efficacy itself held constant. Educational background had three missing cases based on the "N/A" response option; these cases were excluded.

Results

Dropout analyses were conducted using logistic regression with attrition versus completion of all study questionnaires predicted from the study variables of baseline self-efficacy and final grades. These variables did not affect the odds of completing all waves of data collection, suggesting an absence of systematic attrition in regard to tested variables.

Table 2 shows mean self-efficacy and self-efficacy bias at each time point. Three *t*-tests for each variable indicated that both mean self-efficacy and mean self-efficacy bias decreased significantly over time. Table 3 shows mean self-efficacy and self-efficacy bias at each time point for over-efficacious and non over-efficacious students. These figures also show decreasing self-efficacy and self-efficacy bias over time. The proportion of participants in each group at each time point are illustrated in Figure 2, also showing how over-efficaciousness decreased over time.

Table 2

Mean self-efficacy and mean self-efficacy bias

	n	Self-efficacy	Self-efficacy bias
Baseline	197	2.67 (.85) [a]	.99 (1.14)[a]
Mid-semester	197	2.12 (.83)[b]	.44 (1.06)[b]
End-of-semester	124	1.88 (.84) [c]	.15 (.82)[c]

Note. SDs are in parentheses. Positive bias scores indicate over-efficaciousness
a, b, c = pairs of means marked by different superscript letters in each column differed significantly from each other at p<.001

Table 3

Mean self-efficacy and self-efficacy bias for over-efficacious and non over-efficacious participants

Self-efficacy				
	n	Over-efficacious	n	Non over-efficacious
Baseline	133	2.93 (.71)	64	2.13 (.86)
Mid-semester	85	2.53 (.85)	112	1.81 (.67)
End-of-semester	36	2.22 (.78)	88	1.75 (.83)

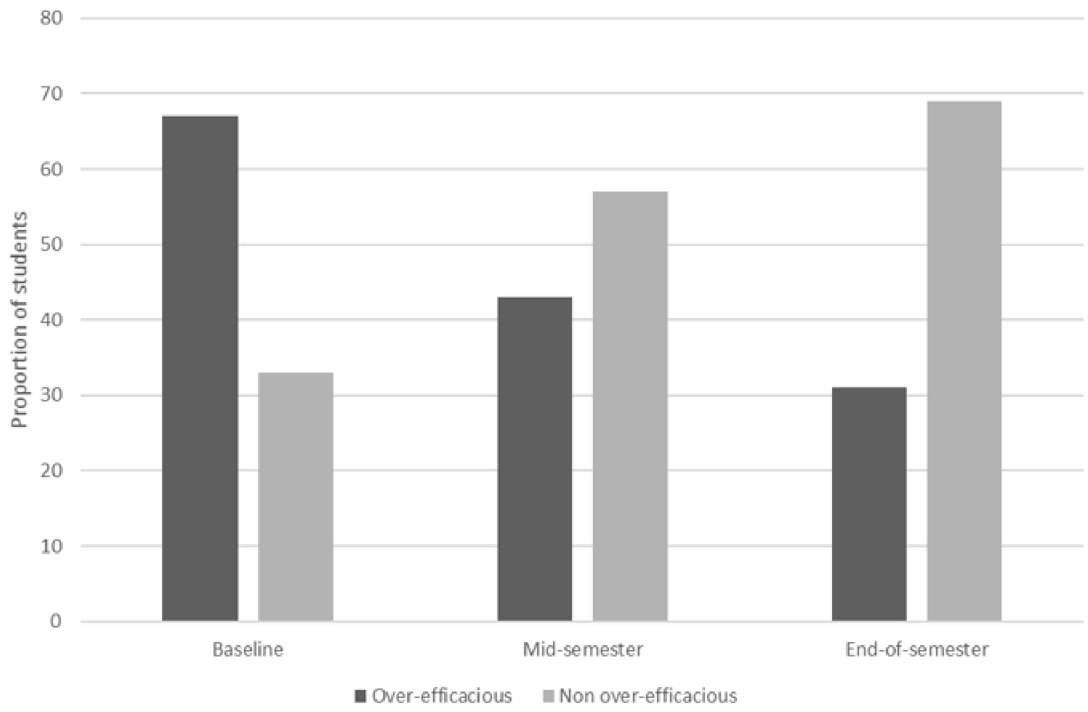


Figure 2. Proportion of over- and non over-efficacious participants at each time point

Results of the three logistic regression analyses are presented in Tables 4-6. The final model for the beginning-of-semester explained a significant amount of variance in the data, $\chi^2(4)=18.04$, $p=.001$, Nagelkerke pseudo $R^2=.12$, and was a good fit to the data, Pearson $\chi^2(189)=194.16$, $p=.38$. Using the model, 71.1% of cases were correctly classified, a slight improvement (3.6%) over the null model (from which we would predict that all participants would be over-efficacious based on observed frequencies only) (see Figure 1). Participants with lower cognitive ability and lower agreeableness had greater odds of being over-efficacious. For each unit decrease in cognitive ability or agreeableness, the odds of being over-efficacious increased by 20% and 12% respectively.

Table 4

Logistic regression: Over-efficacious students compared with peers at the beginning-of-semester

	B(SE)	Odds ratio	95% CI lower	95% CI upper
Intercept	1.88 (1.10)			
Cohort	.49 (.33)	1.62	.86	3.09
Self-efficacy	.057 (.02)*	1.06	1.02	1.10
Cognitive Ability	-.22 (.07)*	.80	.69	.93
Agreeableness	-.13 (.06)*	.88	.77	.99

Note. n=194. * p<.05

The final model for mid-semester explained a significant amount of variance in the data, $\chi^2(4)=36.72$, $p<.001$, Nagelkerke pseudo $R^2=.23$, and was a good fit to the data, Pearson $\chi^2(189)=196.50$, $p=.34$. Using the model, 72.2% of cases were correctly classified, an improvement of 15% over the null model. Significant predictor variables were age and self-esteem. Younger participants had greater odds of being over-efficacious, at a rate of 4% greater odds per year. For each unit increase in self-esteem, the odds of being over-efficacious increased by 13%.

Table 5

Logistic regression: Over-efficacious students compared with peers at mid-semester

	B(SE)	Odds ratio	95% CI lower	95% CI upper
Intercept	-3.84* (.90)			
Cohort	.74 (.33)*	2.10	1.11	3.98
Self-efficacy	.06 (.02)*	1.06	1.02	1.11
Age	-.04 (.02)*	.96	.93	1.00[a]
Self-esteem	.12 (.04)*	1.13	1.04	1.22

Note. n=194; [a] The rounding of odds ratios to two decimal places obscures the 95% CI upper for Age (OR: .998). *= p<.05

The final end-of-semester model also explained a significant amount of variance in the data, $\chi^2(4)=27.63$, $p<.001$, Nagelkerke pseudo $R^2=.29$, and was a good fit to the data, Pearson $\chi^2(117)=139.38$, $p=.08$. Using the model, 78.7% of cases were correctly classified, an improvement of 8.2% over the null model. Significant predictor variables were self-esteem and SES. For each unit increase in self-esteem, participants had 11% greater odds of being over-efficacious. For each unit increase in SES, participants had 1% greater odds of being over-efficacious. While this might seem a trivial amount, we note that the scale score for SES has a range of 275 points, providing substantial scope for the influence of SES.

Table 6

Logistic regression: Over-efficacious students compared with peers at the end-of-semester

	B(SE)	Odds ratio	95% CI lower	95% CI upper
Intercept	-13.65 (3.61)*			
Cohort	.62 (.46)	1.88	.77	4.61
Self-efficacy	.07 (.03)*	1.07	1.01	1.14
Self-esteem	.10 (.05)*	1.11	1.00 [a]	1.22
SES	.01 (<.01)*	1.01	1.00 [a]	1.02

Note. *n*=122; [a] The rounding of odds ratios to two decimal places obscures the 95% CI lower bounds for self-esteem and SES (OR: 1.001). *= *p*<.05

Discussion

The aim of the present study was to identify individual differences and contextual characteristics which differentiate over-efficacious tertiary students from their non over-efficacious peers. Logistic regression models provided support for the influence of both cognitive and motivated self-evaluation processes on over-efficaciousness, as well as differences based on participants' age, personality (agreeableness) and SES. Remaining personality variables, ASE, GSE, and participant sex did not contribute to any of the regression models.

At the beginning of semester, those with lower cognitive ability had greater odds of being over-efficacious. This finding is consistent with an "unskilled and unaware" effect, where lower cognitive ability negatively impacts both the individual's ability to successfully complete tasks, and simultaneously deprives them of insight into their own (lack of) ability (Ehrlinger et al., 2008; Kruger & Dunning, 1999). More antagonistic (less agreeable) individuals also had greater odds of being over-efficacious at baseline. More agreeable individuals are characterised as modest and humble, while more antagonistic individuals may be immodest and proud (McCrae, 1987); for example, agreeableness correlates negatively with self-estimated intelligence (Furnham, Moutafi, & Chamorro-Premuzic, 2005). A feature of antagonism is the tendency to set oneself against others (McCrae, 1987). As such, biases such as the better-than-average effect (Alicke & Govorun, 2005) might have a stronger influence on antagonistic individuals, contributing to over-efficaciousness.

At mid-semester, younger students were more likely to be over-efficacious. This is consistent with previous studies indicating better accuracy of self-evaluations in older students, and less overconfidence in decision-making in older adults (de Bruin et al., 2012; Grimes, 2002). Generational differences may also contribute to this effect, with recent research suggesting that millennials (Generation Y), who comprise the bulk of our sample, are more overconfident than other generations (Twenge, 2014). Students with higher self-esteem also had greater odds of over-efficaciousness at mid-semester. People with high self-esteem are more likely to have unrealistically favourable views about themselves generally (Baumeister et al., 2003) and are also more pre-occupied with self-enhancement than those with low self-esteem (Sedikides & Strube, 1997). In this context, those with high self-esteem are more likely to process and remember information which illuminates positive aspects of the self (e.g., good grades, positive feedback) and to selectively forget, diminish or misattribute information which threatens self-esteem (e.g., poor grades, criticism) (Sedikides & Strube, 1997). Thus, over-efficaciousness in those with high self-esteem may stem from consistent unrealistically positive self-views and motivated self-enhancement processes.

Self-esteem was also a key predictor at end-of-semester, suggestive of an influence which persists in spite of potentially sobering domain-specific feedback received during the course. At end-of-semester, students from higher SES backgrounds also had greater odds of being over-efficacious. In our sample, this finding is likely underscored by higher self-efficacy in higher SES students (r=.145, p=.04) with no corresponding academic performance advantage (r=.01, p=.87). The self-efficacy beliefs of higher SES students may be anchored to societal and parental expectations of successful educational performance, and/or students' own higher aspirations (Cervone & Peake, 1986; Neuenschwander et al., 2007).

Implications

Recent research indicates that beliefs which exceed capacity are associated with poorer academic outcomes (Dunlosky & Rawson, 2012; Talsma, Norris, et al., 2017), and researchers suggest that accurate self-beliefs are most adaptive in terms of academic self-regulation and performance (Boekaerts & Rozendaal, 2010; Stankov & Lee, 2017). The present study identifies some characteristics of individuals likely to be over-efficacious, which may have applications for reducing over-efficaciousness in educational settings, thus potentially curbing some of the associated risks.

In the challenging transition to university, many beginning university students' expectations go unmet early in their studies, and there is growing interest in interventions designed to ease this transition (Brinkworth, McCann, Matthews, & Nordström, 2009). Findings from the present study could be incorporated into such interventions, to help address over-efficaciousness. For example, drawing on findings that younger and higher SES students have greater odds of over-efficaciousness, university information and pathway planning sessions at high SES pre-tertiary institutions could be tailored to provide information regarding the prevalence of overconfidence in first-year university students, and how this relates to performance outcomes. Students may benefit from information about the differences in the university environment and how to manage their expectations and beliefs accordingly (Brinkworth et al., 2009).

Once a student's university career is underway, a key issue underlying the formation of self-efficacy beliefs is feedback (Hattie & Timperley, 2007). An understanding of one's level of mastery of skills or knowledge is the primary basis of self-efficacy beliefs (Bandura, 1997). Thus, a potential way of increasing accuracy of self-efficacy beliefs is through the use of accurate performance feedback (Gist & Mitchell, 1992). For those with high self-esteem, which was found to increase the odds of over-efficaciousness at both mid-semester and end-of-semester, this may create somewhat of a catch-22. Over-efficaciousness could potentially be reduced by more closely attending to negative feedback and taking corrective action, but for individuals with high-self-esteem, the desire to defend self-worth against unwelcome feedback may negate its value to self-regulatory processes which could be informed by more realistic appraisals of performance (Crocker & Park, 2004). Thus, the pursuit or maintenance of high self-esteem may be associated with self-regulatory costs, where short-term gratification for self-worth protection is traded off against long term goals. For example, negative feedback which might prompt adaptive behaviours (e.g., greater effort, a change in strategy) is ignored or downplayed, or externally attributed, because it is uncomfortable or threatening to one’s positive self-image (Crocker, Moeller, & Burson, 2010). This tendency may be pronounced if students’ education to this point has been subject to policies influenced by the self-esteem movement, in which case, criticism, negative feedback and poor grades may be particularly unfamiliar and confronting (Baumeister et al., 2003; Dinham, 2010). Thus, we suggest that the framing of feedback may be important in increasing the accuracy of self-efficacy beliefs of students with high self-esteem.

In framing feedback to increase accuracy of self-efficacy beliefs without activating self-worth protection processes, a focus on communicating key messages about the meaning and purpose of feedback may be beneficial. While shying away from negative feedback and failing grades is argued to be unwarranted and potentially counterproductive (Baumeister et al., 2003; Dinham, 2010), steps may be taken to ensure that students understand that feedback is not an evaluation of them as a person, but of their performance in a specific context (Nicol

& Macfarlane-Dick, 2006). Framing feedback in a constructive way, by identifying opportunities to close the gap between actual performance and desired performance, may also be advantageous (Nicol & Macfarlane-Dick, 2006).

Limitations

Although the final regression models explained significant variance in the data, substantial additional variance remained unexplained, particularly at the beginning-of-semester. Further research using additional predictor variables is warranted. Our non-experimental findings preclude conclusions about causality. Internal consistency for conscientiousness was lower than recommended; it is unknown whether a more reliable measure of conscientiousness would have been a significant predictor. Several relationships underlying our findings were not consistent with some previous literature; e.g., first-generation tertiary students' self-efficacy beliefs were not lower than those of their peers, and higher SES students did not outperform their lower SES counterparts. To the extent that these findings may be idiosyncratic, we recommend that further research be undertaken.

Conclusion

Key factors identified as distinguishing over-efficacious students from their peers include lower cognitive ability and higher self-esteem, providing an indication of the role of both cognitive processes and motivated self-evaluation processes in over-efficaciousness. Low agreeableness, younger age and higher SES also characterised over-efficacious students. These findings may inform theory development and interventions, such as university orientation programs and feedback policies, designed to reduce over-efficaciousness and its associated academic risks.

References

Alicke, M. D., & Govorun, O. (2005). The better-than-average effect. In M. D. Alicke, D. Dunning, & J. Krueger (Eds.), *The self in social judgment* (Vol. 1, pp. 85-106). New York: Psychology Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 5, pp. 307-337). Greenwich, CT: Information Age.

Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological science in the public interest, 4*(1), 1-44. doi:10.1111/1529-1006.01431

Bloom, J. L. (2007). (mis) reading social class in the journey towards college: Youth development in urban america. *Teachers College Record, 109*(2), 343-368.

Boekaerts, M., & Rozendaal, J. S. (2010). Using multiple calibration indices in order to capture the complex picture of what affects students' accuracy of feeling of confidence. *Learning and Instruction, 20*(5), 372-382. doi:10.1016/j.learninstruc.2009.03.002

Brinkworth, R., McCann, B., Matthews, C., & Nordström, K. (2009). First year expectations and experiences: Student and teacher perspectives. *Higher Education*, *58*(2), 157-173. doi:10.1007/s10734-008-9188-3

Brown, J. D., Collins, R. L., & Schmidt, G. W. (1988). Self-esteem and direct versus indirect forms of self-enhancement. *Journal of Personality and Social Psychology, 55*(3), 445. doi:10.1037/0022-3514.55.3.445

Caprara, G. V., Vecchione, M., Alessandri, G., Gerbino, M., & Barbaranelli, C. (2011). The contribution of personality traits and self-efficacy beliefs to academic achievement: A longitudinal study. *British Journal of Educational Psychology, 81*(1), 78-96. doi:10.1348/2044-8279.002004

Carpenter, D. M., & Clayton, G. (2014). Measuring the relationship between self-efficacy and math performance among first-generation collegebound middle school students. *Middle Grades Research Journal, 9*(2), 109-126.

Cervone, D., & Peake, P. K. (1986). Anchoring, efficacy, and action: The influence of judgmental heuristics on self-efficacy judgments and behavior. *Journal of Personality and Social Psychology, 50*(3), 492. doi:10.1037/0022-3514.50.3.492

Chen, P. P. (2003). Exploring the accuracy and predictability of the self-efficacy beliefs of seventh-grade mathematics students. *Learning and Individual Differences, 14*(1), 77-90. doi:10.1016/j.lindif.2003.08.003

Crocker, J., Moeller, S., & Burson, A. (2010). The costly pursuit of self-esteem. In R. H. Hoyle (Ed.), *Handbook of personality and self-regulation* (pp. 403-429). West Sussex, UK: Blackwell Publishing.

Crocker, J., & Park, L. E. (2004). The costly pursuit of self-esteem. *Psychological Bulletin, 130*(3), 392-414. doi:10.1037/0033-2909.130.3.392

Dahl, M., Allwood, C. M., Rennemark, M., & Hagberg, B. (2010). The relation between personality and the realism in confidence judgements in older adults. *European Journal of Ageing, 7*(4), 283-291. doi:10.1007/s10433-010-0164-2

de Bruin, W. B., Parker, A. M., & Fischhoff, B. (2012). Explaining adult age differences in decision-making competence. *Journal of Behavioral Decision Making, 25*(4), 352-360. doi:10.1002/bdm.712

De Feyter, T., Caers, R., Vigna, C., & Berings, D. (2012). Unraveling the impact of the big five personality traits on academic performance: The moderating and mediating effects of self-efficacy and academic motivation. *Learning and Individual Differences, 22*(4), 439-448. doi:10.1016/j.lindif.2012.03.013

Dinham, S. (2010). Self-esteem: Caution-do not over-inflate [online]. *Teacher: The National Education Magazine,* (Aug 2010), 6-11.

Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-ipip scales: Tiny-yet-effective measures of the big five factors of personality. *Psychological Assessment, 18*(2), 192-203. doi:10.1037/1040-3590.18.2.192

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction, 22*(4), 271-280. doi:10.1016/j.learninstruc.2011.08.003

Dunning, D. (2005). *Self-insight: Roadblocks and detours on the path to knowing thyself*. New York: Psychology Press.

Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes, 105*(1), 98-121. doi:10.1016/j.obhdp.2007.05.002

Ehrlinger, J., Mitchum, A. L., & Dweck, C. S. (2016). Understanding overconfidence: Theories of intelligence, preferential attention, and distorted self-assessment. *Journal of Experimental Social Psychology, 63*, 94-100. doi:10.1016/j.jesp.2015.11.001

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). *Manual for kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Furnham, A., Moutafi, J., & Chamorro-Premuzic, T. (2005). Personality and intelligence: Gender, the big five, self-estimated and psychometric intelligence. *International*

Journal of Selection and Assessment, 13(1), 11-24. doi:10.1111/j.0965-075X.2005.00296.x

Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management Review, 17*(2), 183-211. doi:10.5465/AMR.1992.4279530

Grimes, P. W. (2002). The overconfident principles of economics student: An examination of a metacognitive skill. *The Journal of Economic Education, 33*(1), 15-30. doi:10.1080/00220480209596121

Hacker, D. J., Bol, L., & Keener, M. C. (2008). Metacognition in education: A focus on calibration. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 429-455). New York: Psychology Press.

Inman, W. E., & Mayes, L. (1999). The importance of being first: Unique characteristics of first generation community college students. *Community College Review, 26*(4), 3-22. doi:doi:10.1177/009155219902600402

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121-1134. doi:10.1037/0022-3514.77.6.1121

Lee, C., & Bobko, P. (1994). Self-efficacy beliefs: Comparison of five measures. *Journal of Applied Psychology, 79*(3), 364. doi:10.1037/0021-9010.79.3.364

Marsh, H. W. (1990). The structure of academic self-concept: The marsh/shavelson model. *Journal of Educational Psychology, 82*(4), 623-636. doi:10.1037/0022-0663.82.4.623

McCrae, R. R. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*(1), 81-90. doi:10.1037/0022-3514.52.1.81

Neuenschwander, M. P., Vida, M., Garrett, J. L., & Eccles, J. S. (2007). Parents' expectations and students' achievement in two western nations. *International Journal of Behavioral Development, 31*(6), 594-602. doi:10.1177/0165025407080589

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218. doi:10.1080/03075070600572090

Pallier, G., Wilkinson, R., Danthiir, V., Kleitman, S., Knezevic, G., Stankov, L., & Roberts, R. D. (2002). The role of individual differences in the accuracy of confidence judgments. *Journal of General Psychology, 129*(3), 257-299. doi:10.1080/00221300209602099

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (mslq)*. Ann Arbor, MI: The University of Michigan.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353-387. doi:10.1037/a0026838

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Schwarzer, R., & Jerusalem, M. (1995). Generalized self-efficacy scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35-37). Windsor, UK: NFER-NELSON.

Sedikides, C., & Strube, M. J. (1997). Self-evaluation: To thine own self be good, to thine own self be sure, to thine own self be true, and to thine own self be better. *Advances in experimental social psychology, 29*, 209-269. doi:10.1016/S0065-2601(08)60018-0

Sheldrake, R. (2016). Differential predictors of under-confidence and over-confidence for mathematics and science students in England. *Learning and Individual Differences, 49*, 305-313. doi:10.1016/j.lindif.2016.05.009

Sheldrake, R., Mujtaba, T., & Reiss, M. J. (2014). Calibration of self-evaluations of mathematical ability for students in England aged 13 and 15, and their intentions to study non-compulsory mathematics after age 16. *International Journal of Educational Research, 64*, 49-61. doi:10.1016/j.ijer.2013.10.008

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*(3), 417-453. doi:10.3102/00346543075003417

Stankov, L., & Lee, J. (2014). Overconfidence across world regions. *Journal of Cross-Cultural Psychology, 45*(5), 821-837. doi:10.1177/0022022114527345

Stankov, L., & Lee, J. (2017). Self-beliefs: Strong correlates of mathematics achievement and intelligence. *Intelligence, 61*, 11-16. doi:10.1016/j.intell.2016.12.001

Talsma, K., Norris, K., & Schuz, B. (2017). Miscalibration of self-efficacy and academic performance: Self-efficacy ≠ self-fulfilling prophecy. *Manuscript submitted for publication*.

Talsma, K., Schüz, B., Schwarzer, R., & Norris, K. (2017). I believe, therefore I achieve (and vice versa): A meta-analytic cross-lagged panel analysis of self-efficacy and academic performance. *Manuscript submitted for publication*.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*(2), 193. doi:10.1037/0033-2909.103.2.193

Twenge, J. M. (2014). *Generation me-revised and updated: Why today's young americans are more confident, assertive, entitled--and more miserable than ever before*. New York: Simon and Schuster.

Chapter 6: Discussion

The overall aim of the present thesis was to develop a deeper understanding of the dynamics of the relationship between self-efficacy and performance in educational settings. While a substantial body of extant literature considers the relationship between self-efficacy and academic performance (e.g., Honicke & Broadbent, 2016; Multon et al., 1991; Richardson et al., 2012; Robbins et al., 2004), the present thesis aimed to move beyond the conventional focus on the positive association between the two variables, typically approached from the self-efficacy → academic performance perspective, and instead consider complexities in the relationship which have received limited research attention. Specifically, questions of reciprocity and calibration between self-efficacy and academic performance were addressed.

The question of reciprocity reflects the chicken-and-egg conundrum in the relationship between self-efficacy and academic performance: is it *I believe therefore I achieve*, or vice versa, or both? The question has long been theoretically resolved in favour of a reciprocal relationship, but has lacked convincing empirical support (Chapter 2, Study 1; Chapter 3, Study 2). The question of calibration of self-efficacy beliefs related to the accuracy of university students' academic self-efficacy beliefs and how calibration for one academic outcome (the deviation of self-efficacy beliefs from academic performance) related to future academic performance (Chapter 4, Study 3). Competing hypotheses about the optimum level of self-efficacy in terms of academic outcomes (accuracy of self-efficacy beliefs, over-efficaciousness, or under-efficaciousness) were considered. Characteristics which may differentiate over-efficacious students from their peers were also considered (Chapter 5, Study 3).

The key aims, findings and conclusions for the four studies comprising this thesis are outlined in Table 1. This discussion chapter will summarise and integrate the main findings from the included studies, highlight implications for theory, research and educational practice stemming from this research, discuss potential limitations and directions for future research in this area, and draw overall conclusions.

Table 1

Summary of aims, findings and conclusions of the present thesis

Chapter	Aim	Findings	Conclusions
2: I believe, therefore I achieve (and vice versa): A meta-analytic cross-lagged panel analysis of self-efficacy and academic performance	• To determine whether self-efficacy and academic performance influence each other reciprocally over time, and to compare the strength of directional effects. • To explore the moderating effect of individual differences and methodological factors in the self-efficacy ↔ academic performance relationship. • To approach the question using meta-analytic cross-lagged panel analysis, overcoming the limitations of previous unidirectional research.	• Self-efficacy and academic performance were reciprocally related; self-efficacy and performance had unique positive effects on each other. • The effect of performance on subsequent self-efficacy was significantly stronger than the reverse. • Reciprocity was not evident in children. • Reciprocal relationships were stronger where studies used methodologies consistent with self-efficacy theory.	• Findings support the conceptualisation of the self-efficacy ↔ performance relationship as a feedback loop. • Recommendations regarding methodological approaches in self-efficacy research appear to have a genuine impact on research quality. • Applied strategies which leverage the synergies in the bidirectional relationship may be more successful than considering either self-efficacy or performance in isolation.
3: Reciprocity in the self-efficacy ↔ academic performance relationship: The effect of measurement order	• To determine whether the strength of directional effects in the reciprocal self-efficacy ↔ academic performance relationship is influenced by measurement order (i.e., whether self-efficacy is measured prior or subsequent to performance at a single measurement point). • To compare the strength of reciprocal effects between studies with reversed measurement orders.	• Measurement order moderated the reciprocal self-efficacy ↔ performance relationship. • For studies which measured performance first at each measurement wave, a meta-analytic cross-lagged panel analysis indicated that self-efficacy was a stronger predictor of performance than the reverse (an opposite finding to Chapter 1, in which self-efficacy was measured first at each measurement wave).	• Findings provide further support for the self-efficacy ↔ performance feedback loop. • Self-efficacy is a stronger predictor of subsequent performance when participants have performance exposure prior to making initial self-efficacy judgements. • This may reflect greater accuracy, or calibration, of self-efficacy beliefs based on task experience.
4: Miscalibration of self-efficacy and academic performance: Self-efficacy ≠ self-fulfilling prophecy	• To explore the calibration of students' self-efficacy beliefs with their academic performance, focusing on task types and levels of specificity not previously considered in the literature. • To determine whether self-efficacy calibration (accuracy of self-efficacy, over-efficaciousness, or under-efficaciousness) predicts future performance outcomes.	• Participants were generally poorly calibrated, with under-efficaciousness prevalent for written assignments and exams, and over-efficacious common for subject grades. • Generally, weaker performers were over-efficacious and stronger performers were under-efficacious. • Stronger performance on similar subsequent tasks was predicted by self-efficacy which tended towards under-efficaciousness.	• Inaccurate self-efficacy beliefs are prevalent in learners, suggesting that the self-efficacy ↔ performance relationship is not a closed loop. • The direction of miscalibration varies depending on the specificity/proximity of the academic outcome. • Under-efficaciousness was associated with stronger performance. • Over-efficaciousness may be a risk factor for poor performance.
5: When belief exceeds capacity: Which factors differentiate over-efficacious students from their peers?	• To determine whether individual differences (age, sex, cognitive ability, self-beliefs, personality traits) and contextual factors (SES, educational background) differentiate over-efficacious students from their peers, at three time points over an academic semester.	• At the beginning of semester, over-efficacious students were lower in cognitive ability and agreeableness. • At mid-semester, over-efficacious students were younger with higher self-esteem. • At the end of semester, over-efficacious students had higher self-esteem and SES.	• Key results suggest an initial influence of cognitive processes and a sustained influence of motivated self-evaluation processes on miscalibrated self-efficacy beliefs. • Findings may inform educational policies and interventions for students likely to be over-efficacious, reducing academic risks.

Reciprocity of self-efficacy and academic performance

In chapters 2 and 3, reciprocity in the relationship between self-efficacy and academic performance was considered. While social cognitive theory locates self-efficacy within a model of reciprocal determinism (Bandura, 1997), limited empirical evidence has directly examined this supposition, and little is known about the comparative strength of directional effects in the relationship. This is largely because of the methodological and analytical approaches of the extant research. Previous studies which bear on this issue have used unidirectional approaches, or measurements staggered over time, limiting the conclusions which could be drawn. To address this, chapters 2 and 3 applied the technique of cross-lagged panel analysis to meta-analytic data from relevant studies (Figure 1, Chapter 2). Analysing studies using panel data in which both self-efficacy and performance were measured simultaneously on two measurement waves provided a more rigorous examination of reciprocity than has been previously conducted.

Chapters 2 and 3 provide an important contribution to the debate regarding reciprocity between self-efficacy and academic performance, enabling the isolation of the unique effects of self-efficacy and performance on each other over time, and the identification of variables which moderate the relationship. The rigorous systematic approach and use of meta-analytic panel analysis in the testing of hypotheses are key strengths of these studies. Findings from both studies provide support for the social cognitive theory model of reciprocal determinism between person factors and behavioural factors. The finding of a unique impact of performance on subsequent self-efficacy aligns with the proposition that performance experiences reflect mastery, which is a key source of self-efficacy judgements (Bandura, 1997). The magnitude of this effect in comparison to that of self-efficacy on performance provides some support for researchers who have suggested that the impact of performance on self-efficacy has been underemphasised in the literature (e.g., Vancouver et al., 2001). The unique impact of self-efficacy on subsequent performance is consistent with conceptions of self-efficacy as a mobilising influence on academic performance (e.g., Pajares, 1996, 2006). This finding also casts doubt on suggestions that self-efficacy simply reflects past performance (Heggestad & Kanfer, 2005).

Chapter 2 included studies in which self-efficacy was measured prior to performance at each measurement wave, while Chapter 3 included studies in which self-efficacy was measured subsequent to performance at each measurement wave. It was anticipated that the dynamics of the feedback loop may manifest differently when self-efficacy was measured after performance, because of the use of newly gleaned performance-related information in the formation of self-efficacy beliefs (Gist & Mitchell, 1992).

When performance was measured prior to self-efficacy at time 1 (Chapter 3), self-efficacy at time 1 had a significantly stronger relationship with performance at time 2 than it did when self-efficacy was measured prior to performance at time 1 (Chapter 2). The reverse also applied: when self-efficacy was measured prior to performance at time 1 (Chapter 2), performance at time 1 had a significantly stronger relationship with self-efficacy at time 2 than it did when performance was measured prior to self-efficacy at time 1 (Chapter 3).

This difference may be explained by the effect of task experience on the accuracy of self-efficacy beliefs: when students had an opportunity to perform before making self-efficacy judgements, these judgements were more accurate and thus more strongly predicted subsequent performance (Gist & Mitchell, 1992). A comparison of cross-lagged effects between chapters 2 and 3 suggest that through exposure to performance-feedback cycles, students' self-efficacy and academic performance are likely to become more calibrated with each other over time.

These studies also explored the ways in which reciprocity between self-efficacy and academic performance varies as a function of moderator variables (sample characteristics and methodological approaches).

Moderator analyses indicated that the self-efficacy ↔ performance relationship varied as a function of participant age. Specifically, the relationship was reciprocal for adults, but in children, performance predicted subsequent self-efficacy, but the path from self-efficacy to performance was not significant. This may reflect the greater metacognitive capacity of adult learners, or the differences between the learning environments of adults and children (for example, adult learning tasks may be more challenging, feedback may be more specific).

The reciprocal relationship also varied based on the lag time between measurements, with shorter intervals associated with stronger relationships, as suggested by self-efficacy theory. The degree of match between self-efficacy and performance operationalisations also moderated the relationship, with a match in terms of the specificity of the construct being measured showing stronger relationships than less matched constructs. Again, this is consistent with the recommendations of self-efficacy theorists. Finally, the relationship was moderated by the type of scale used to measure self-efficacy. Self-efficacy studies using Likert scales generally found smaller effects than those using the unidirectional scale recommended by self-efficacy theorists (e.g., confidence in ability from 0 to 100%). Overall, these findings suggest that cross-lagged effects were stronger where studies used methodologies consistent with recommendations of self-efficacy theorists (e.g., Bandura, 2006, 2012).

Together, chapters 2 and 3 contribute to the literature by providing support for reciprocal determinism between self-efficacy and academic performance, and by providing important insights into the circumstances under which the strength of bidirectional effects may vary in the self-efficacy ↔ performance relationship.

Implications for theory and research. Chapters 2 and 3 provided support for reciprocal determinism as posited by self-efficacy theory (Bandura, 1977). When controlling for other paths in the models, both self-efficacy and performance uniquely predicted each other over time, regardless of whether, on any measurement occasion, self-efficacy was measured first (chapter 2) or performance was measured first (chapter 3). Findings thus support the suppositions that self-efficacy has a mobilising effect on performance, and that performance experiences, which reflect mastery, are used as a source of self-efficacy beliefs. Moreover, a comparison of the cross-lagged paths in chapters 2 and 3 provided support for more detailed theoretical conceptualisations of the relationship between self-efficacy and performance, for example, theories highlighting the effects of task analysis on self-efficacy beliefs (Gist & Mitchell, 1992). Effect sizes were modest, consistent with previous meta-analyses. This is a point I return to later.

Future self-efficacy research will be strengthened by approaches that take the reciprocal nature of the relationship with performance into account. Research which considers the effect of self-efficacy on performance without controlling for previous performance will ascribe a greater role to self-efficacy than may be warranted. Similarly, research which considers the influence of performance experiences or mastery on self-efficacy beliefs without accounting for previous self-efficacy may over-estimate the effect of mastery on self-efficacy. Results suggest that the greatest power to detect effects in self-efficacy research will be observed by researchers who follow methodological guidelines for self-efficacy research, including matching the self-efficacy and performance constructs and using unidirectional rather than Likert scales. Naturally, the timeframe over which data is collected will be informed by the specific research question, but shorter intervals appear likely to show stronger effects.

The research implications stemming from chapters 2 and 3 can be summed up in a simple phrase: everything matters. The findings of this thesis suggest do not imply that any particular method is superior in all circumstances, but that even small differences in approach (e.g., scale selection, measurement order, timing of measurement, participant age) have a measurable effect on research outcomes, and that informed decisions about how to approach self-efficacy research questions should be made accordingly. In the same vein, when considering synthesising findings across studies, all of these differences (and still more that were not considered here) are likely to account for heterogeneity in groups of included studies.

While stability of self-efficacy and performance over time was not an issue considered directly in chapters 2 and 3, the relevant paths in the cross-lagged path analyses suggest that the strongest relationships are the autocorrelations: self-efficacy at time 1 is the best predictor of self-efficacy and time 2, and the same trend applies to performance (effect sizes in the range β= 0.5–0.6, compared with β =0.1–0.2 for cross-lagged paths). Thus, while self-efficacy is more accurate in predicting performance if beliefs include more information on previous performance, as shown in Chapter 3, students appear to rely more on their previous self-efficacy beliefs than performance information when forming subsequent self-efficacy beliefs. This suggests that opportunities exist to further explore how interventions might enhance the function of the feedback loop, leading to better calibration of self-efficacy and performance over time. Having said that, previous research aimed at improving accuracy of self-beliefs has met with limited success, across interventions including feedback, practice, and incentives for accuracy (Hacker, Bol, & Keener, 2008). Key theorists maintain that self-efficacy is not a trait, partly because the strength of self-efficacy beliefs varies across domains (Bandura, 2006). However, if the accuracy of self-efficacy proves to be impervious to interventions designed to improve it, further exploration of a potential trait component of self-efficacy may be fruitful. This exploration may be guided by research approaches investigating a stable confidence factor for performance postdictions (Stankov et al., 2015) and a stable higher order construct which subsumes self-efficacy, along with self-esteem, neuroticism and locus of control (Judge et al., 2002).

Practical implications. In a setting where strong self-efficacy beliefs appear to be increasingly considered as an end in themselves (Ritchie, 2015), it behooves educators and education policy makers to consider that the relationship between self-efficacy and academic performance is more complicated than it might appear at first glance. No matter how many meta-analyses show a positive relationship between self-efficacy and academic performance, caution still needs to be exercised in interpreting correlational data. Thus, while a substantial body of evidence (including that presented here) indicates that, overall, higher self-efficacy beliefs are associated with stronger academic performance, analyses at this level potentially obscure important complexities in the relationship.

It appears self-evident that neither self-efficacy beliefs nor performance capacity develop in a vacuum. Focusing on both in concert is more likely to take advantage of the synergies inherent in the feedback loop which appears to underlie the reciprocal relationship. In addition, as the feedback loop continues to function, it is suggested that performance-feedback cycles will increase the accuracy of self-efficacy beliefs, decreasing the gap between self-efficacy beliefs and performance outcomes which underscores miscalibration. Thus, considering the findings of the present dissertation as a whole, feedback presents itself as a key issue for educators and education policy makers (Dinham, 2010; Hattie & Timperley, 2007). Sensitively framed but accurate feedback may improve calibration of self-efficacy, particularly in over-efficacious students. Self-monitoring activities, such as keeping a log of academic tasks along with feedback and reflections (Nicol & Macfarlane-Dick, 2006), may be incorporated into the curriculum in order to help develop the metacognitive skills thought to underlie accurate calibration. These approaches may be tailored to reach those students more likely to show over-efficaciousness, such as younger students, those from high SES backgrounds, or those with high self-esteem, as outlined in chapter 5. The purpose of such an approach is not necessarily to curb self-efficacy beliefs overall, but to provide an indication of where resources could be most gainfully invested.

Calibration of self-efficacy and academic performance

Chapters 4 and 5 focused on calibration (accuracy of self-efficacy beliefs) and miscalibration (over- or under-efficaciousness) in university students. Chapter 4 (study 3) examined the impact of self-efficacy calibration on future performance outcomes – thus addressing the ongoing controversy in the self-efficacy literature of what constitutes the optimum level of self-efficacy in terms of strong outcomes. Having identified over-efficaciousness as a risk factor for poor performance, chapter 5 (study 4) considered person and contextual variables which differentiated over-efficacious students from their accurate and under-efficacious peers.

Some previous research has considered self-efficacy beliefs for mathematics tasks, mostly in children and adolescents (e.g., Chen & Zimmerman, 2007; Chen, 2003), but little is known about the calibration of university students' self-efficacy beliefs for their course assessments, or how self-efficacy calibration relates to future academic performance. This latter question is of particular interest given contradictory positions in the literature regarding whether accuracy of self-beliefs (e.g., Artino, 2012; Stankov & Lee, 2017), over-efficaciousness (e.g., Bandura, 1997; Pajares, 2006), or under-efficaciousness (e.g., Stone, 1994; Vancouver & Kendall, 2006) should be associated with better subsequent performance.

Another important element of the present studies was that the performance outcomes were measured in naturalistic learning settings; that is, they formed part of students' assessments and were subject to naturally occurring timeframes and pressures. This is in contrast to the previous research on calibration of self-efficacy which has largely focused on mathematics tasks designed for the purpose of the study and conducted in a single session.

In Chapter 4 (study 3), participants' self-efficacy judgements were compared to their performance on two written assignments and two exams, and to their subject grades overall. Self-efficacy accuracy (the absolute difference between self-efficacy and performance) and bias (the signed difference; over- and under-efficaciousness) were computed and used to predict subsequent performance. Students were found to be largely inaccurate in their self-efficacy judgements. Under-efficaciousness was commonly observed for written assignments and exams. This is in contrast to previous findings regarding academic self-efficacy, which tend to show that students are more likely to be over-efficacious than under-efficacious (e.g., Chen & Zimmerman, 2007; Chen, 2003). This finding also diverges from the extensive literature documenting a tendency towards unrealistically positive self-evaluations across a range of domains and tasks (see Moore & Healy, 2008; Stone, 1994; Zabrucky, 2010). This tendency was demonstrated in this study when considering subject level outcomes, with over-efficaciousness pronounced for subject grades.

While both task-level and subject-level self-efficacy judgements were made in the absence of relevant mastery information (given that students were first-year undergraduates) it is interesting to note that different judgements overall characterised task- and subject-level outcomes. A possible explanation is that the differences in the timing and concreteness of the outcome play a key role here: for task-based assessments, self-efficacy judgements were made in the days leading up to submission, whereas for subject-level assessments, self-efficacy judgements were made at the beginning of semester, rendering the outcome temporally more distal than the task-based judgements. In a context where a specific type of performance is imminent, self-efficacy judgements are more likely subject to the sobering influences of anxiety, defensive pessimism, and bracing for the worst. In contrast, for more abstract and distal outcomes, judgements may be more subject to self-serving biases and increased optimism.

Another key trend was that self-efficacy exceeded performance capacity for low-achievers, while high-achievers' self-efficacy beliefs tended to underestimate their capacity to perform. For low achievers, this is consistent with previous findings regarding the Dunning-Kruger effect or "unskilled and unaware" effect (Kruger & Dunning, 1999), where a lack of ability manifests not only in poor performance on the task, but also in a lack of awareness of what would constitute strong performance on the task. As discussed in chapter 4, high achievers may be more subject to defensive pessimism, imposter phenomenon, or simply more cautious self-evaluations.

A key contribution of the chapter 4 study was to predict future performance using a measure of bias from a previous analogous task. Students who were under-efficacious with regard to their first assessments performed better on their second assessments, for both written assignments and exams. Findings suggest self-efficacy is not a self-fulfilling prophecy: that is, some students who believe they can, objectively cannot, while some students who believe they cannot, evidently can. The theoretical and practical implications of these findings are discussed further below.

Chapter 4 provided the impetus for exploring variables which might characterise over-efficacious students, who are potentially at risk of poor academic outcomes. Thus, in chapter 5 (study 4), I examined whether individual differences and contextual factors could be used to differentiate over-efficacious students from their peers. Individual-level variables considered were age and sex, as well as cognitive ability, more global-level self-beliefs including self-esteem, self-concept, and generalised self-efficacy, and the Five Factor Model (FFM) personality traits. Environmental variables included SES and educational background.

Self-efficacy scores at the beginning, middle and end of semester were compared to final subject grades and participants were identified as either over-efficacious or non over-efficacious at each time point. Compared to their peers, over-efficacious students had greater odds of having lower cognitive ability and lower agreeableness at the beginning of semester. At mid-semester, over-efficacious students had greater odds of being younger and having higher self-esteem. High self-esteem also characterised those students who were over-efficacious at the end of semester. Students from higher SES backgrounds were also more likely to be over-efficacious. Discussion focused mainly on self-esteem protection as a potential sustaining mechanism for over-efficaciousness, in that a tendency towards self-enhancement and selective processing of favourable information may also support unrealistically positive self-efficacy beliefs (e.g., Crocker, Moeller, & Burson, 2010). Chapter 5 provides an important preliminary understanding of the personal and environmental variables which accompany over-efficaciousness in university students.

Together, chapters 4 and 5 provide important insights into how calibration of self-efficacy beliefs in university students varies according to the type of task, the specificity of the performance outcome, and the ability level of the student, as well building our understanding of how calibration relates to future performance and of which characteristics differentiate over-efficacious students from their peers.

Implications for theory and research. Findings from chapter 4 did not support the proposition that self-efficacy which exceeds capacity to perform (over-efficaciousness) is associated with stronger performance outcomes, nor do the findings support the proposed benefit of accuracy of self-efficacy beliefs on future performance. In this study, self-efficacy beliefs which fell short of capacity were associated with better subsequent performance. As such, findings are inconsistent with self-efficacy theory on face value, and are more consistent with explanations of self-efficacy which focus on discrepancy reduction processes

(Carver & Scheier, 1982; Powers, 1973). That is, under-efficacious students may, for example, work harder and perform because of the misperception that they are not as capable as they actually are, whereas over-efficacious students, perceiving no lack of ability, experience no discrepancy in need of addressing and thus refrain from allocating needed resources to study. Having said that, intermediary variables such as goal-setting, effort and persistence were not measured. It is possible that over-efficacious students set higher goals and allocated more resources to the tasks in question, as would be predicted by self-efficacy theory, but that this did not result in improved performance. Perhaps, instead, over-efficacious students misjudged the demands of the task or their own mastery of subject material, and consequently misapplied these additional resources.

With the above caveat in mind, the present research has some bearing on research conducted exploring the potential negative effects of self-efficacy on performance. Work by Vancouver and colleagues (Vancouver, Thompson, Tischner, & Putka, 2002; Vancouver et al., 2001) has shown a negative effect of self-efficacy on subsequent performance, including in learning settings (Vancouver & Kendall, 2006). In this study, high self-efficacy negatively impacted academic resource allocation and subsequent exam performance (Vancouver & Kendall, 2006). Findings are broadly consistent with this previous research, with the qualification that it is not high self-efficacy in itself, but self-efficacy which exceeds performance capacity which appears to negatively impact on academic performance. Under-efficacious students, in contrast, regardless of their actual level of self-efficacy, potentially perform more strongly because of the motivation provided by the belief that they are less capable than they actually are.

Future theorising may further specify relevant models to account for observed complexities such as reciprocity and calibration in the relationship between self-efficacy and academic performance, along with other context-dependent differences in the dynamics of the relationship identified elsewhere including feedback (Beattie, Woodman, Fakehy, & Dempsey, 2016), the experience of failure (Hardy III, 2014; Schmidt & DeShon, 2009), and the availability of rewards (Stirin Tzur, Ganzach, & Pazy, 2016). The role of cognitive and motivated processes in the formation of miscalibrated self-beliefs discussed in chapter 5 are also worthy of further study and potential incorporation into theoretical models of the relationship between self-efficacy and academic performance. A more detailed and flexible model may help to account for the equivocal findings in the literature to date – in terms of positive, negative or null effects of self-efficacy on performance.

Our findings also have some bearing on discussions regarding the strength of the association between self-efficacy and academic performance. The effect size of the relationship between self-efficacy and performance in previous meta-analyses, as well as in the meta-analytic findings of the present thesis, is small-to-medium in magnitude. Similar cross-sectional effect sizes were reported in chapter 3. Educational contexts are complex and many additional variables may account for the modest nature of these effects; however, a parsimonious explanation is that the miscalibration of self-efficacy attenuates the strength of the relationship. In chapter 4, under-efficacious students tended to be stronger performers, and vice versa. For this reason, the linear relationship between self-efficacy and performance is flatter than it would be if all students were accurately calibrated. The relationship remains positive, because many under-efficacious students' self-efficacy beliefs, though lower than warranted by objective outcomes, are still higher than those of many over-efficacious students, who perform more poorly. Using self-efficacy calibration for the first written report, from chapter 4, as an example, Figure 1 highlights this potential explanation. As shown in the figure, the linear relationship between actual self-efficacy and academic performance is affected by over-efficaciousness at the lower end of the performance scale, and by under-efficaciousness at the higher end of the performance scale. This is reflected in a flatter linear relationship, and therefore a smaller correlation, between self-efficacy and performance.

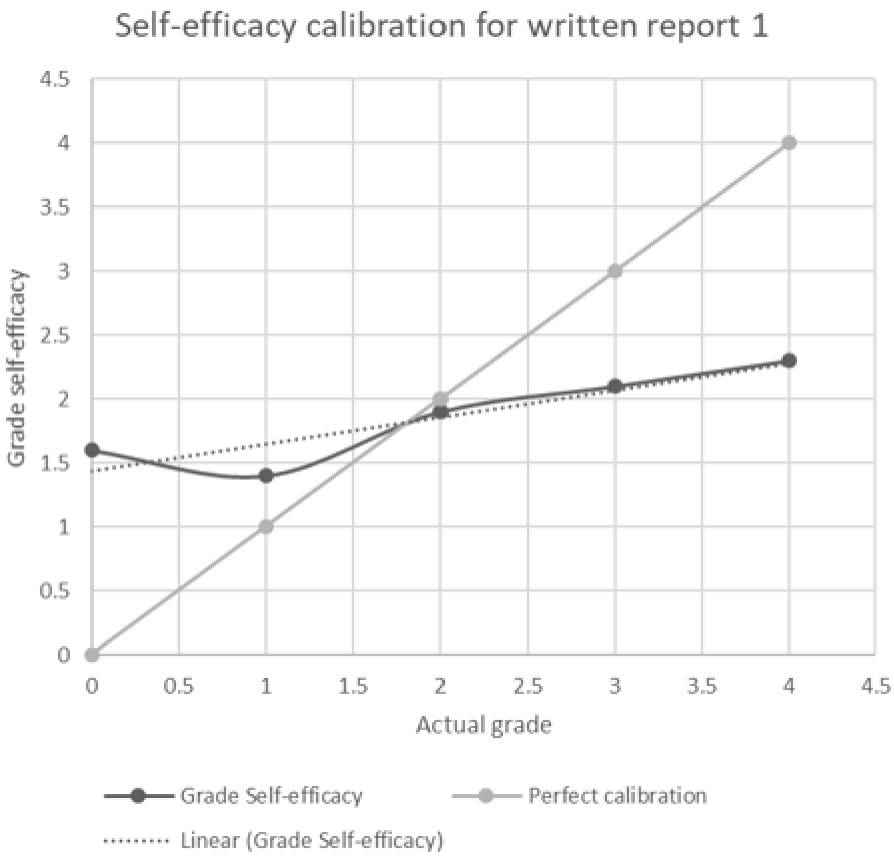


Figure 1. Attenuation of the correlation between self-efficacy and performance based on miscalibration

Practical implications. Conventional notions of the relationship between self-efficacy and performance in educational settings provide us with two main images of learners: one who has high self-efficacy and performs well, and one who has low self-efficacy and performs poorly – and whose self-efficacy beliefs should thus be bolstered (e.g., Carman, 2015; Corno, 1982; Sewell & St George, 2009). This thesis' findings regarding calibration of self-efficacy with objective academic performance suggest that this conventional notion oversimplifies the relationship between self-efficacy and academic performance, and that further conceptualisations, addressing the issue of calibration, are necessary. To this end, Figure 2 is provided as a suggested taxonomy of academic self-efficacy beliefs which includes both conventional conceptions, in which self-efficacy beliefs are calibrated with performance

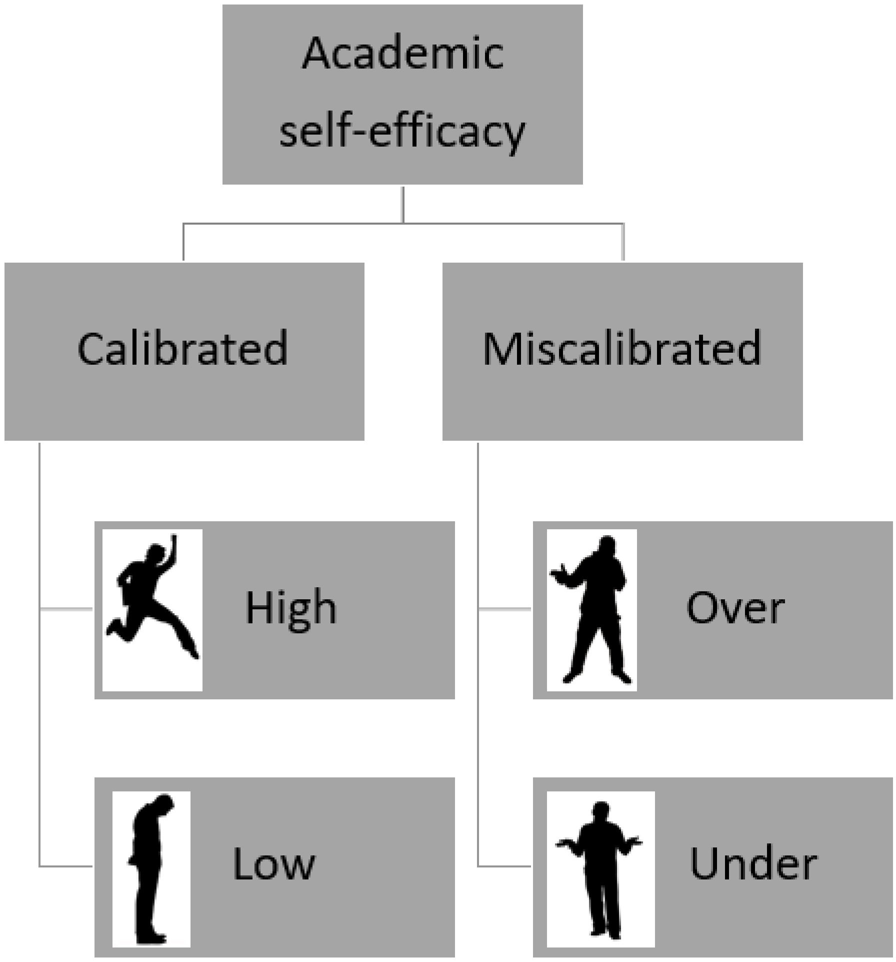


Figure 2. A taxonomy of self-efficacy beliefs

capacity, and viewpoints based on miscalibration of self-efficacy beliefs. The conventional view of students with high self-efficacy who perform well and students with low self-efficacy who perform poorly is elaborated by including over-efficacious students, who have stronger self-efficacy beliefs than are borne out in their performance outcomes, and under-efficacious students, whose performance capacity exceeds their belief in their ability. Because the present findings suggest that miscalibration of self-efficacy beliefs may be quite prevalent in adult learners, it appears to be an oversimplification to assume that low self-efficacy beliefs should always be bolstered, or that all students with strong self-efficacy beliefs will perform well. The findings from the present study may encourage educators to closely observe the dynamics of self-efficacy and performance in their own students, and may inspire them to broach the topic of self-efficacy policies in their schools with a view to developing approaches which more closely reflect the complexities of the relationship.

A search of “academic self-efficacy” on Google indicates that “how to improve self-efficacy in students” is one of the most commonly used search strings on the topic, yielding some 2 million relevant pages. The websites of adult education providers, ranging from American and Australian universities (Garlin, 2014; Kirk, n. d.; Nelson & Cooper, 2014), to global conglomerates like education.com, opencolleges.com.au and pearsoned.com all provide resources designed to improve self-efficacy (Briggs, 2014; Haskell, 2016; Mayer, 2010). Whole books are devoted to teaching educators how to foster self-efficacy beliefs in university students (Ritchie, 2015). Ultimately, however, the findings of the present thesis suggest that applied approaches such as these, which appear to rest on a simplistic interpretation of older literature (i.e., high self-efficacy is always good; self-efficacy should always be bolstered), are not necessarily going to be of benefit to all students in all circumstances. While suggestions to intentionally induce mildly negative expectations in order to improve performance are not unprecedented (e.g., Stone, 1994), further research is needed to weigh up the relative effects of such approaches on performance, engagement and wellbeing measures. Based on the present findings, however, I echo the caution that efforts to bolster self-efficacy in isolation from the development of relevant knowledge and skills may be counterproductive (Vancouver & Kendall, 2006). A more circumspect viewpoint than provided by conventional notions appears to be warranted: a strong sense of self-efficacy may be motivating in some circumstances, but self-efficacy which exceeds capacity to perform may not always be beneficial, and self-efficacy which falls short of capacity perform may not always be detrimental. In particular, findings suggest that increasing the accuracy of self-efficacy beliefs for over-efficacious students may be of benefit, because over-efficaciousness was associated with poorer performance, whereas trying to boost the self-efficacy beliefs of under-efficacious students may be a disservice to them, given that under-efficaciousness most strongly predicted good subsequent performance in the present results.

Limitations and directions for future research

In this section I begin with more specific limitations of the present research, before moving onto broader limitations and directions for future research.

In the meta-analyses in chapters 2 and 3, the use of strict inclusion criteria was necessitated by the cross-lagged panel approach. For this reason, from the many hundreds of studies identified in the initial literature searches, only a small number were ultimately included. There are a number of reasons, as discussed within, to believe that the findings presented are valid in spite of these small sample sizes. Nonetheless, further replication is necessary before strong conclusions are drawn.

As is relatively common with psychology research, there was a lack of diversity in samples across all studies, in that there was limited representation of non-western cultures, learning disabled participants, and individuals from different social strata. Further research with culturally and socially diverse samples may provide additional support for the present findings, or highlight further complexities in the relationship between self-efficacy and academic performance.

In the primary studies of the present thesis, participants were adult learners only. Meanwhile, the meta-analytic results indicated that the reciprocal relationship between self-efficacy and academic performance was moderated by age. Thus, there is reason to suspect that the findings regarding miscalibration of self-efficacy and academic performance, as well as findings regarding the characteristics of over-efficacious students, may not hold for children. There may be other miscalibration patterns specific to school-aged children, and characteristics which differ from those discussed here may increase the odds of msiscalibration in this age group.

A key limitation in the present studies is that intermediary influences such as effort and persistence were not measured. As mentioned above, it may be that over-efficacious students exerted more effort as self-efficacy theory would predict, but that this did not translate into stronger performance because it was misapplied – perhaps because of a misunderstanding of specific areas of weakness, for example. Research using expanded models which consider preparatory behaviours is necessary to determine where exactly the potential disconnect between self-efficacy and academic performance occurs. Thus, further research considering these intervening variables is clearly warranted.

Another fruitful avenue for further research is studies investigating whether interventions can lead to more accurate self-efficacy beliefs, and whether this has flow-on effects to academic performance. A critical area for future research in this area is provision of feedback. Interventions focusing on feedback and rewards have recently been considered in sports psychology (Beattie et al., 2016), and it was found that the type of feedback moderated the relationship between self-efficacy and performance (i.e., with detailed feedback, self-efficacy was positively related to subsequent performance, whereas with minimal feedback, self-efficacy was negatively related to subsequent performance). A similar trend may be observed in future research in educational settings.

This observation also raises a question with regard to the present analyses. I have speculated that accurate feedback may increase accuracy of calibration and that this may in turn result in better self-regulation and academic outcomes. An issue in the present thesis is that detailed feedback was provided to students in line with institutional assessment policy, but it is unknown to what extent students engaged with this feedback and used it in the formation of subsequent self-efficacy beliefs. Differential engagement with feedback was not explored or controlled in the present studies; thus, potential impacts of this are not considered. Similarly, while detailed task outlines and assessment rubrics were provided to all students in advance of task performance, it is unknown to what degree students engaged with these materials. Over- and under-efficaciousness may rest in part on differing degrees of engagement with task information provided to students, which was not measured.

Identification of further characteristics which differentiate over- and under-efficacious students may also have benefits for application in educational settings. The present study included person characteristics (e.g., personality, self-esteem) and environmental characteristics (e.g., SES, educational background), whereas variables at the junction between the individual and the subject area specifically are also likely to be influential. For example, in research exploring calibration of self-concept, subject interest, anxiety, perceived utility of the subject area, and relative difficulty of the subject differentially predicted over- and under-confidence (Sheldrake, 2016). Institutional factors such as perceived competence of teachers and availability of academic support may also play a role.

Qualitative analysis involving interviews with participants identified as over- or under-efficacious may also provide a deeper understanding of the cognitive and motivational processes underlying miscalibrated self-efficacy judgements. For example, as suggested by Vancouver and colleagues (2006; 2002; 2001) and discussed above, it is possible that under-efficaciousness predicts better subsequent performance because of processes discussed in control and discrepancy theories, whereby more effort is exerted to close the perceived gap between belief and reality. It is also possible, however, that the self-efficacy beliefs of under-efficacious individuals are not actually a reflection of genuine doubt regarding capacity to achieve, but a strategy used (intentionally or otherwise) by high performers, for example, to fire themselves up to work harder, or to protect themselves from less-than-perfect outcomes.

Overall, in spite of the considerable body of literature already addressing the relationship between self-efficacy and academic performance, it appears that many questions remain and that future research will continue to identify subtleties in the relationship.

Context specificity issues

Before concluding, a word is warranted regarding the ways in which the particularities of the academic context potentially impact on the dynamics of the relationship between self-efficacy and performance (e.g., Schunk, 1996).

An issue in educational research generally is the challenge of measuring learning accurately; typically, performance assessments are used as proxies for learning (Crisp, 2012) and the benefits and drawbacks of different approaches to assessment are the subject of much debate (Boud & Falchikov, 2006). Assessments may not provide the validity and precision required of a psychometrically sound measure of student learning if assessments are, for example, poorly designed or unreliably graded (Boud & Falchikov, 2007; Boud, Lawson, & Thompson, 2013). This means that interpretations of the findings of research using assessment performance as a measure of learning are valid to the extent that assessment is itself a valid indicator of student learning. This is relevant in the context of the present findings, where self-efficacy beliefs are expressed in terms of ability to achieve particular performance (i.e., assessment) outcomes. It may be that students are able to accurately judge what they have learned, but are less able to accurately judge how they will perform, because of a lack of alignment between material to be learned and assessment of that material.

In addition, with regard to the measurement of self-efficacy, distinctions have been drawn between contexts in which the required skills and/or knowledge are already mastered, and those in which the required skills and/or knowledge are yet to be learned. The former context has been referred to as self-efficacy for performance, with the latter referred to as self-efficacy for learning (Schunk, 1996), or preparatory self-efficacy (Bandura & Locke, 2003).

It has been suggested that educational settings are likely to be characterised by self-efficacy for learning, or preparatory self-efficacy (Vancouver & Kendall, 2006). In contrast, consider the self-efficacy-for-performance context in which self-efficacy theory was first developed: snake phobia (Bandura & Adams, 1977). If the participant was required to, for example, touch a snake, there is no ambiguity with regard to the required behaviour, nor is there any lack of actual skill or knowledge that would prevent this behaviour. Self-efficacy for, e.g., health behaviours operates in a context with many of the same features: while participants may be reluctant to exercise regularly or quit smoking, the required behaviour is clear and, in principle at least, everyone is fundamentally capable of performing the behaviour. In these circumstances, self-efficacy beliefs are likely to be accurate, and thus a strong predictor of the behaviour in question.

In educational settings, in contrast, self-efficacy judgements are often made under conditions of uncertainty and ambiguity, both with regard to the precise features and demands of the behaviour in question, and with regard to the degree to which the individual has already mastered the skills or knowledge required to perform the behaviour. The issue is further complicated by the timing dynamics at play – while the necessary skills may not be mastered right now, they may have been by the time the performance of the behaviour is required, which may be some way off.

It has been pointed out that these complexities deem self-efficacy judgements prone to influences from heuristics and biases (Stone, 1994), and to greater risk of miscalibration and potential negative effects of self-efficacy on subsequent performance (Vancouver & Kendall, 2006; Verhaeren, 2012). Put simply, in authentic learning settings, there are far more reasons for self-efficacy beliefs to be inaccurate, and thus less predictive of future performance, when compared to other types of settings.

I note, however, that there are many studies conducted regarding academic self-efficacy which seek to narrowly define self-efficacy and isolate behaviours in time – perhaps for the very purpose of removing much of the complexity that would typically characterise self-efficacy judgements made in authentic learning settings. This can be seen in approaches which, for example, ask participants to indicate their degree of confidence to answer a particular mathematics question shown to them (e.g., Chen, 2003). These studies can be more accurately conceived of as explorations of self-efficacy for performance. These differences within academic paradigms present something of a conundrum, because, while the above mathematics example may be more likely to result in more accurate self-efficacy beliefs, the intermediary processes such as goal-setting, effort, and persistence would also be less likely to come into play compared to more naturalistic educational contexts (e.g., as performance is imminent, there is no particular occasion for extra investment of effort). Both research approaches are of value, and in combination they may strike a balance between tight study specification and the exploration of authentic learning behaviours.

On the whole, research designs may be strengthened by taking into consideration the potential differences in the dynamics of the relationship between self-efficacy and academic performance, depending on the context of the behaviour and the nature of the types of tasks being considered. These differences may be evident both across and within domains, and may influence the generalisibility and synthesisability of findings regarding self-efficacy and performance.

Conclusion

Taken together, the findings of this thesis contribute to a deeper understanding of the dynamics and subtleties underlying the relationship between self-efficacy and academic performance. Conventional perspectives of the relationship between self-efficacy and academic performance generally hold that self-efficacy is the antecedent in the relationship, and that strong self-efficacy beliefs are a desirable outcome in learning settings because they lead to better performance. With a pedigree reaching as far back as The Little Engine that Could, the *I believe therefore I achieve* perspective of self-efficacy and performance has strong intuitive appeal. While these maxims may apply in some contexts, the findings of the present thesis point to the need for a more nuanced perspective of the role of self-efficacy in academic performance. Specifically, self-efficacy does not act on educational outcomes in one direction; the findings of chapters 2 and 3 indicate that bidirectional influences should be considered in research and practice. While the relationship appears to be reciprocal, this does not necessarily mean that self-efficacy beliefs accurately reflect performance capacity. Chapter 4 shows that, in university students, miscalibration appears to be prevalent, manifesting as either over- or under-efficaciousness depending on the domain specificity and temporal proximity of the outcome. Over-efficaciousness appears to present the greatest risk of negative performance outcomes, and chapter 5 indicates that there are observable person and contextual variables which may differentiate over-efficacious students from their peers. It is hoped that these understandings will prompt further research and theory development, and the consideration of more finely-tuned educational policies which acknowledge that a one-size-fits-all approach to self-efficacy is unlikely to be of universal benefit.

References

Artino, A. R. (2012). Academic self-efficacy: From educational theory to instructional practice. *Perspectives on Medical Education, 1*(2), 76-85. doi:10.1007/s40037-012-0012-5

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191. doi:10.1037/0033-295X.84.2.191

Bandura, A. (1978). The self system in reciprocal determinism. *American Psychologist, 33*(4), 344-358. doi:10.1037/0003-066X.33.4.344

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall, Inc.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 5, pp. 307-337). Greenwich, CT: Information Age.

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management, 38*(1), 9-44. doi:10.1177/0149206311410606

Bandura, A., & Adams, N. E. (1977). Analysis of self-efficacy theory of behavioral change. *Cognitive Therapy and Research, 1*(4), 287-310. doi:10.1007/BF01663995

Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology, 88*(1), 87-99. doi:10.1037/0021-9010.88.1.87

Beattie, S., Woodman, T., Fakehy, M., & Dempsey, C. (2016). The role of performance feedback on the self-efficacy–performance relationship. *Sport, Exercise, and Performance Psychology, 5*(1), 1-13. doi:10.1037/spy0000051

Beaudoin, M., & Desrichard, O. (2011). Are memory self-efficacy and memory performance related? A meta-analysis. *Psychological Bulletin, 137*(2), 211-241. doi:10.1080/13803390802232667

Berrington, A., Smith, P., & Sturgis, P. (2006). *An overview of methods for the analysis of panel data*. Retrieved from eprints.ncrm.ac.uk/415/1/MethodsReviewPaperNCRM-007.pdf

Boekaerts, M., & Rozendaal, J. S. (2010). Using multiple calibration indices in order to capture the complex picture of what affects students' accuracy of feeling of confidence. *Learning and Instruction, 20*(5), 372-382. doi:10.1016/j.learninstruc.2009.03.002

Bol, L., & Hacker, D. J. (2012). Calibration research: Where do we go from here? *Frontiers in Psychology, 3,* 229: 221-226. doi:10.3389/fpsyg.2012.00229

Bong, M. (2001). Role of self-efficacy and task-value in predicting college students' course performance and future enrollment intentions. *Contemporary Educational Psychology, 26*(4), 553-570. doi:10.1006/ceps.2000.1048

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*(1), 1-40. doi:10.1023/A:1021302408382

Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education, 31*(4), 399-413. doi:10.1080/02602930600679050

Boud, D., & Falchikov, N. (2007). *Rethinking assessment in higher education: Learning for the longer term*. New York: Routledge.

Boud, D., Lawson, R., & Thompson, D. G. (2013). Does student engagement in self-assessment calibrate their judgement over time? *Assessment & Evaluation in Higher Education, 38*(8), 941-956. doi:10.1080/02602938.2013.769198

Briggs, S. (2014). *Why self-esteem hurts learning but self-confidence does the opposite.* Retrieved from http://www.opencolleges.edu.au/informed/features/self-efficacy-and-learning/

Brown, S. D., Tramayne, S., Hoxha, D., Telander, K., Fan, X., & Lent, R. W. (2008). Social cognitive predictors of college students' academic performance and persistence: A meta-analytic path analysis. *Journal of Vocational Behavior, 72*(3), 298-308. doi:10.1016/j.jvb.2007.09.003

Caprara, G. V., Vecchione, M., Alessandri, G., Gerbino, M., & Barbaranelli, C. (2011a). The contribution of personality traits and self-efficacy beliefs to academic achievement: A longitudinal study. *British Journal of Educational Psychology, 81*(1), 78-96. doi:10.1348/2044-8279.002004

Caprara, G. V., Vecchione, M., Alessandri, G., Gerbino, M., & Barbaranelli, C. (2011b). The contribution of personality traits and self-efficacy beliefs to academic achievement: A longitudinal study. *British Journal of Educational Psychology, 81*(1), 78-96. doi:10.1348/2044-8279.002004

Carman, L. K. (2015). *Low-achieving students' perspectives regarding their experiences as recipients of instructional support delivered in a general education setting.* The State University of New Jersey-New Brunswick. ProQuest Dissertations database. (3738650)

Carver, C. S., & Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality-social, clinical, and health psychology. *Psychological Bulletin, 92*(1), 111-135. doi:10.1037/0033-2909.92.1.111

Chae, S. E. (2012). *Contributions of prior knowledge, motivation, and strategies to korean college students' l2 writing development.* University of Maryland. PsycINFO database. (AAI3495368)

Chen, P., & Zimmerman, B. (2007). A cross-national comparison study on the accuracy of self-efficacy beliefs of middle-school mathematics students. *The Journal of Experimental Education, 75*(3), 221-244. doi:10.3200/JEXE.75.3.221-244

Chen, P. P. (2003). Exploring the accuracy and predictability of the self-efficacy beliefs of seventh-grade mathematics students. *Learning and Individual Differences, 14*(1), 77-90. doi:10.1016/j.lindif.2003.08.003

Corno, L. (1982). *Where there's a way there's a will: Self-regulating the low-achieving student*. Paper presented at the 66th Annual Meeting of the American Educational Research Association, New York.

Crisp, G. T. (2012). Integrative assessment: Reframing assessment practice for current and future learning. *Assessment & Evaluation in Higher Education, 37*(1), 33-43. doi:10.1080/02602938.2010.494234

Crocker, J., Moeller, S., & Burson, A. (2010). The costly pursuit of self-esteem. In R. H. Hoyle (Ed.), *Handbook of personality and self-regulation* (pp. 403-429). West Sussex, UK: Blackwell Publishing.

Dinham, S. (2010). Self-esteem: Caution-do not over-inflate [online]. *Teacher: The National Education Magazine, August 2010*, 6-11.

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction, 22*(4), 271-280. doi:10.1016/j.learninstruc.2011.08.003

Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes, 105*(1), 98-121. doi:10.1016/j.obhdp.2007.05.002

Feltz, D. L., Chow, G. M., & Hepler, T. J. (2008). Path analysis of self-efficacy and diving performance revisited. *Journal of Sport and Exercise Psychology, 30*(3), 401-411. doi:10.1123/jsep.30.3.401

Foster, N. L., Was, C. A., Dunlosky, J., & Isaacson, R. M. (2016). Even after thirteen class exams, students are still overconfident: The role of memory for past exam performance in student predictions. *Metacognition and Learning*, 1-19. doi:10.1007/s11409-016-9158-6

Garlin, F. (2014). *Identifying opportunities to enhance students' academic self-efficacy.* Retrieved from https://www.westernsydney.edu.au/__data/assets/pdf_file/0005/711473/Enhancing_students_academic_SE_workshop.pdf

Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management Review, 17*(2), 183-211. doi:10.5465/AMR.1992.4279530

Hacker, D. J., Bol, L., & Bahbahani, K. (2008). Explaining calibration accuracy in classroom contexts: The effects of incentives, reflection, and explanatory style. *Metacognition and Learning, 3*(2), 101-121. doi:10.1007/s11409-008-9021-5

Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology, 92*(1), 160-170. doi:10.1037/0022-0663.92.1.160

Hacker, D. J., Bol, L., & Keener, M. C. (2008). Metacognition in education: A focus on calibration. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 429-455). New York: Psychology Press.

Hardy III, J. H. (2014). Dynamics in the self-efficacy–performance relationship following failure. *Personality and Individual Differences, 71*, 151-158. doi:10.1016/j.paid.2014.07.034

Haskell, N. (2016). *Classroom strategies to improve student self-efficacy and learning outcomes.* Retrieved from http://www.pearsoned.com/education-blog/improve-self-efficacy-learning-outcomes/

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. doi:10.3102/003465430298487

Heggestad, E. D., & Kanfer, R. (2005). The predictive validity of self-efficacy in training performance: Little more than past performance. *Journal of Experimental Psychology: Applied, 11*(2), 84-97. doi:10.1037/1076-898X.11.2.84

Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review, 17*, 63-84. doi:10.1016/j.edurev.2015.11.002

Hwang, M. H., Choi, H. C., Lee, A., Culver, J. D., & Hutchison, B. (2016). The relationship between self-efficacy and academic achievement: A 5-year panel analysis. *Asia-Pacific Education Researcher, 25*(1), 89-98. doi:10.1007/s40299-015-0236-3

Judge, T. A., & Bono, J. E. (2001). Relationship of core self-evaluations traits – self-esteem, generalized self-efficacy, locus of control, and emotional stability – with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology, 86*(1), 80-92. doi:10.1037//0021-9010.86.1.80

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? *Journal of Personality and Social Psychology*, *83*(3), 693-710. doi:10.1037//0022-3514.83.3.693

Kenney-Benson, G. A., Pomerantz, E. M., Ryan, A. M., & Patrick, H. (2006). Sex differences in math performance: The role of children's approach to schoolwork. *Developmental Psychology, 42*(1), 11-26. doi:10.1037/0012-1649.42.1.11

Kirk, K. (n. d.). *Self-efficacy: Helping students believe in themselves.* Retrieved from https://serc.carleton.edu/NAGTWorkshops/affective/efficacy.html

Klassen, R. M. (2004). A cross-cultural investigation of the efficacy beliefs of south asian immigrant and anglo canadian nonimmigrant early adolescents. *Journal of Educational Psychology, 96*(4), 731-742. doi:10.1037/0022-0663.96.4.731

Klassen, R. M., & Usher, E. L. (2010) Self-efficacy in educational settings: Recent research and emerging directions. *Advances in Motivation and Achievement: Vol. 16 Part A* (pp. 1-33). Bingley, UK: Emerald Group.

Kock, N. (2014). Advanced mediating effects tests, multi-group analyses, and measurement model assessments in pls-based sem. *International Journal of e-Collaboration, 10*(1), 1-13. doi:10.4018/ijec.2014010101

Komarraju, M., & Nadler, D. (2013). Self-efficacy and academic achievement: Why do implicit beliefs, goals, and effort regulation matter? *Learning and Individual Differences, 25*, 67-72. doi:10.1016/j.lindif.2013.01.005

Kontopantelis, E., Springate, D. A., & Reeves, D. (2013). A re-analysis of the cochrane library data: The dangers of unobserved heterogeneity in meta-analyses. *PLoS ONE, 8*(7), e69930. doi:10.1371/journal.pone.0069930

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121-1134. doi:10.1037/0022-3514.77.6.1121

Lindsley, D. H., Brass, D. J., & Thomas, J. B. (1995). Efficacy-performing spirals: A multilevel perspective. *Academy of Management Review, 20*(3), 645-678. doi:10.5465/amr.1995.9508080333

Mayer, R. E. (2010). *Motivation based on self-efficacy*. Retrieved from https://www.education.com/reference/article/motivation-based-self-efficacy/

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52*(1), 81-90. doi:10.1037/0022-3514.52.1.81

Menard, S. (2007). *Handbook of longitudinal research: Design, measurement, and analysis*. New York: Academic Press.

Mitchell, T. R., Hopper, H., Daniels, D., George-Falvy, J., & James, L. R. (1994). Predicting self-efficacy and performance during skill acquisition. *Journal of Applied Psychology, 79*(4), 506-517. doi:10.1037/0021-9010.79.4.506

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review, 115*(2), 502-517. doi:10.1037/0033-295X.115.2.502

Moriarty, B. (2014). Research design and the predictive power of measures of self-efficacy. *Issues in Educational Research, 24*(1), 55-66.

Moritz, S. E., Feltz, D. L., Fahrbach, K. R., & Mack, D. E. (2000). The relation of self-efficacy measures to sport performance: A meta-analytic review. *Research Quarterly for Exercise and Sport, 71*(3), 280-294. doi:10.1080/02701367.2000.10608908

Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38*(1), 30-38. doi:10.1037/0022-0167.38.1.30

Muthen, L. K., & Muthen, B. O. (1998-2012). *Mplus user's guide. Seventh edition.* Los Angeles, CA: Muthen & Muthen.

Nelson, D. B., & Cooper, N. C. (2014). *Academic self-efficacy: The cornerstone of freshmen success.* Retrieved from http://www.nacada.ksu.edu/Resources/Clearinghouse/View-Articles/Academic-self-efficacy-The-cornerstone-of-freshmen-success.aspx

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218. doi:10.1080/03075070600572090

Nietfeld, J. L., Cao, L., & Osborne, J. W. (2006). The effect of distributed monitoring exercises and feedback on performance, monitoring accuracy, and self-efficacy. *Metacognition and Learning, 1*(2), 159-179. doi:10.1007/s10409-006-9595-6

Nunnally, J. C., & Bernstein, I. H. (1994). The assessment of reliability. In I. H. Bernstein & J. C. Nunnally (Eds.), *Psychometric theory* (Vol. 3, pp. 248-292). New York: McGraw-Hill.

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543-578. doi:10.3102/00346543066004543

Pajares, F. (1997). Current directions in self-efficacy research. In M. L. Maehr & P. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 1-49). Greenwich, CT: JAI Press.

Pajares, F. (2002). Overview of social cognitive theory and of self-efficacy. Retrieved from https://www.uky.edu/~eushe2/Pajares/eff.html, 30 June 2017

Pajares, F. (2006). Self-efficacy during childhood and adolescence: Implications for teachers and parents. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 339-368). Greenwich, CT: Information Age.

Pajares, F., & Kranzler, J. (1995). Self-efficacy beliefs and general mental ability in mathematical problem-solving. *Contemporary Educational Psychology, 20*(4), 426-443. doi:10.1006/ceps.1995.1029

Pajares, F., & Miller, M. D. (1995). Mathematics self-efficacy and mathematics performances: The need for specificity of assessment. *Journal of Counseling Psychology, 42*(2), 190-198. doi:10.1037/0022-0167.42.2.190

Pajares, F., & Schunk, D. H. (2001). Self-beliefs and school success: Self-efficacy, self-concept, and school achievement. In R. Riding & S. Rayner (Eds.), *Perception* (pp. 239-266). London: Ablex Publishing.

Pajares, F., & Usher, E. L. (2008). Self-efficacy, motivation, and achievement in school from the perspective of reciprocal determinism. In M. L. Maehr, T. C. Urdan, & S. Karabenick (Eds.), *Advances in motivation and achievement* (Vol. 15, pp. 391-423).

Phan, H. P. (2012). The development of english and mathematics self-efficacy: A latent growth curve analysis. *Journal of Educational Research, 105*(3), 196-209. doi:10.1080/00220671.2011.552132

Powers, W. T. (1973). *Behavior: The control of perception*: Aldine Chicago.

Richard, E. M., Diefendorff, J. M., & Martin, J. H. (2006). Revisiting the within-person self-efficacy and performance relation. *Human Performance, 19*(1), 67-87. doi:10.1207/s15327043hup1901_4

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353-387. doi:10.1037/a0026838

Ritchie, L. (2015). *Fostering self-efficacy in higher education students*. London, UK: Palgrave Macmillan.

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*(2), 261-288. doi:10.1037/0033-2909.130.2.261

Rosenberg, M. (1979). *Conceiving the self*. New York: Basic Books.

Schmidt, A. M., & DeShon, R. P. (2009). Prior performance and goal progress as moderators of the relationship between self-efficacy and performance. *Human Performance, 22*(3), 191-203. doi:10.1080/08959280902970377

Scholz, U., Doña, B. G., Sud, S., & Schwarzer, R. (2002). Is general self-efficacy a universal construct? Psychometric findings from 25 countries. *European Journal of Psychological Assessment, 18*(3), 242-251. doi:10.1027//1015-5759.18.3.242

Schraw, G. (2009). Measuring metacognitive judgments. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 415-429). New York: Routledge.

Schünemann, N., Spörer, N., & Brunstein, J. C. (2013). Integrating self-regulation in whole-class reciprocal teaching: A moderator-mediator analysis of incremental effects on fifth graders' reading comprehension. *Contemporary Educational Psychology, 38*(4), 289-305. doi:10.1016/j.cedpsych.2013.06.002

Schunk, D. H. (1996). *Self-efficacy for learning and performance*. Paper presented at the Annual Conference of the American Educational Research Association, New York, NY.

Schunk, D. H., & Pajares, F. (2004). Self-efficacy in education revisited. In D. M. McInerney & S. Van Etten (Eds.), *Big theories revisited* (Vol. 4, pp. 115-138). Greenwich, CT: Information Age Publishing.

Schwarzer, R., & Jerusalem, M. (1995). Generalized self-efficacy scale. In J. Weinman, S. Wright, & M. Johnston (Eds.), *Measures in health psychology: A user's portfolio. Causal and control beliefs* (pp. 35-37). Windsor, UK: NFER-NELSON.

Schwarzer, R., & McAuley, E. (2016). The world is confounded: A comment on williams and rhodes (2016). *Health Psychology Review, 10*(2), 133-135. doi:10.1080/17437199.2016.1162667

Selig, J. P., & Little, T. D. (2012). Autoregressive and cross-lagged panel analysis for longitudinal data. In B. Laursen, T. D. Little, & N. A. Card (Eds.), *Handbook of developmental research methods* (pp. 265-278). New York: Guilford Press.

Sewell, A., & St George, A. (2009). Developing efficacy beliefs in the classroom. *The Journal of Educational Enquiry, 1*(2).

Shea, C. M., & Howell, J. M. (2000). Efficacy-performance spirals: An empirical test. *Journal of Management, 26*(4), 791-812. doi:10.1177/014920630002600409

Sheldrake, R. (2016). Differential predictors of under-confidence and over-confidence for mathematics and science students in england. *Learning and Individual Differences, 49*, 305-313. doi:10.1016/j.lindif.2016.05.009

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2015). Official user-guide to the p-curve. Retrieved from www.p-curve.com/guide.pdf,

Sitkin, S. B. (1992). Learning through failure: The strategy of small losses. In B. M. Straw & L. L. Cummings (Eds.), *Research in organizational behavior* (Vol. 14, pp. 231-266). Greenwich, CT: JAI Press.

Stajkovic, A. D., & Luthans, F. (1998). Self-efficacy and work-related performance: A meta-analysis. *Psychological Bulletin, 124*(2), 240-261. doi:10.1037/0033-2909.124.2.240

Stankov, L., Kleitman, S., & Jackson, S. (2015). Measures of the trait of confidence. In G. Boyle, D. Saklofske, & G. Matthews (Eds.), *Measures of personality and social psychological constructs* (pp. 158-189). London, UK: Academic Press.

Stankov, L., & Lee, J. (2008). Confidence and cognitive test performance. *Journal of Educational Psychology, 100*(4), 961-976. doi:10.1037/a0012546

Stankov, L., & Lee, J. (2014). Overconfidence across world regions. *Journal of Cross-Cultural Psychology, 45*(5), 821-837. doi:10.1177/0022022114527345

Stankov, L., & Lee, J. (2017). Self-beliefs: Strong correlates of mathematics achievement and intelligence. *Intelligence, 61*, 11-16. doi:10.1016/j.intell.2016.12.001

Stankov, L., Lee, J., Luo, W., & Hogan, D. J. (2012). Confidence: A better predictor of academic achievement than self-efficacy, self-concept and anxiety? *Learning and Individual Differences*, *22*(6), 747-758. doi:10.1016/j.lindif.2012.05.013

Stankov, L., Lee, J., & Paek, I. (2009). Realism of confidence judgments. *European Journal of Psychological Assessment, 25*(2), 123-130. doi:10.1027/1015-5759.25.2.123

Stirin Tzur, K., Ganzach, Y., & Pazy, A. (2016). On the positive and negative effects of self-efficacy on performance: Reward as a moderator. *Human Performance, 29*(5), 362-377. doi:10.1080/08959285.2016.1192631

Stolp, S., & Zabrucky, K. M. (2009). Contributions of metacognitive and self-regulated learning theories to investigations of calibration of comprehension. *International Electronic Journal of Elementary Education, 2*(1), 7-31.

Stone, D. N. (1994). Overconfidence in initial self-efficacy judgments: Effects on decision processes and performance. *Organizational Behavior and Human Decision Processes, 59*(3), 452-474. doi:10.1006/obhd.1994.1069

Stone, N. J. (2000). Exploring the relationship between calibration and self-regulated learning. *Educational Psychology Review, 12*(4), 437-475. doi:10.1023/A:1009084430926

Talsma, K., Schüz, B., Schwarzer, R., & Norris, K. (2017). I believe, therefore i achieve (and vice versa): A meta-analytic cross-lagged panel analysis of self-efficacy and academic performance. *Manuscript submitted for publication*.

Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational psychologist, 39*(2), 111-133. doi:10.1207/s15326985ep3902_3

Vancouver, J. B., & Kendall, L. N. (2006). When self-efficacy negatively relates to motivation and performance in a learning context. *Journal of Applied Psychology, 91*(5), 1146-1153. doi:10.1037/0021-9010.91.5.1146

Vancouver, J. B., Thompson, C. M., Tischner, E. C., & Putka, D. J. (2002). Two studies examining the negative effect of self-efficacy on performance. *Journal of Applied Psychology, 87*(3), 506-516. doi:10.1037/0021-9010.87.3.506

Vancouver, J. B., Thompson, C. M., & Williams, A. A. (2001). The changing signs in the relationships among self-efficacy, personal goals, and performance. *Journal of Applied Psychology, 86*(4), 605. doi:10.1037/0021-9010.86.4.605

Verhaeren, T. (2012). Is a strong sense of self-efficacy always beneficial? *Bulletin of the Transilvania University of Braşov, Series VII: Social Sciences and Law*(1), 193-200.

Viechtbauer, W. (2010). Conducting meta-analyses in r with the metafor package. *Journal of Statistical Software, 36*(3), 1-48. doi:10.18637/jss.v036.i03

Zabrucky, K. M. (2010). Knowing what we know and do not know: Educational and real world implications. *Procedia: Social and Behavioral Sciences, 2*(2), 1266-1269. doi:10.1016/j.sbspro.2010.03.185

Zimmerman, B. J. (1995). Self-efficacy and educational development. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 202-231). Cambridge, UK: Cambridge University Press.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, M. Zeidner, & P. Pintrich (Eds.), *Handbook of self-regulation* (pp. 13-39). California: Academic Press.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into practice, 41*(2), 64-70. doi:10.1207/s15430421tip4102_2

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-motivation for academic attainment: The role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal, 29*(3), 663-676. doi:10.2307/1163261

Appendices

Appendix 1 Ethics approval

Social Science Ethics Officer
Private Bag 01 Hobart
Tasmania 7001 Australia
Tel: (03) 6226 2763
Fax: (03) 6226 7148
Katherine.Shaw@utas.edu.au

HUMAN RESEARCH ETHICS COMMITTEE (TASMANIA) NETWORK

UTAS

28 May 2014

Dr Kimberley Norris
Psychology
Private Bag 30

Student Researcher: Kate Talsma

Sent via email

Dear Dr Norris

Re: MINIMAL RISK ETHICS APPLICATION APPROVAL
Ethics Ref: **H0014107 - Academic performance, mood, and the development of academic self-efficacy**

We are pleased to advise that acting on a mandate from the Tasmania Social Sciences HREC, the Chair of the committee considered and approved the above project on 27 May 2014.

This approval constitutes ethical clearance by the Tasmania Social Sciences Human Research Ethics Committee. The decision and authority to commence the associated research may be dependent on factors beyond the remit of the ethics review process. For example, your research may need ethics clearance from other organisations or review by your research governance coordinator or Head of Department. It is your responsibility to find out if the approval of other bodies or authorities is required. It is recommended that the proposed research should not commence until you have satisfied these requirements.

Please note that this approval is for four years and is conditional upon receipt of an annual Progress Report. Ethics approval for this project will lapse if a Progress Report is not submitted.

The following conditions apply to this approval. Failure to abide by these conditions may result in suspension or discontinuation of approval.

1. It is the responsibility of the Chief Investigator to ensure that all investigators are aware of the terms of approval, to ensure the project is conducted as approved by the Ethics Committee, and to notify the Committee if any investigators are added to, or cease involvement with, the project.

2. Complaints: If any complaints are received or ethical issues arise during the course of the project, investigators should advise the Executive Officer of the Ethics Committee on 03 6226 7479 or human.ethics@utas.edu.au.

3. Incidents or adverse effects: Investigators should notify the Ethics Committee immediately of any serious or unexpected adverse effects on participants or unforeseen events affecting the ethical acceptability of the project.

4. Amendments to Project: Modifications to the project must not proceed until approval is obtained from the Ethics Committee. Please submit an Amendment Form (available on our website) to notify the Ethics Committee of the proposed modifications.

5. Annual Report: Continued approval for this project is dependent on the submission of a Progress Report by the anniversary date of your approval. You will be sent a courtesy reminder closer to this date. **Failure to submit a Progress Report will mean that ethics approval for this project will lapse.**

6. Final Report: A Final Report and a copy of any published material arising from the project, either in full or abstract, must be provided at the end of the project.

Yours sincerely

Katherine Shaw
Executive Officer
Tasmania Social Sciences HREC

Appendix 2.1 Literature search: Database search terms

Scopus

TITLE-ABS-KEY (self-efficacy OR "self efficacy") AND TITLE-ABS-KEY (reciproc* OR longitud* OR "reverse caus*" OR circular OR lag* OR panel OR "growth curve" OR bidirect* OR bi-direct*) AND TITLE-ABS-KEY (academi* OR universit* OR school* OR educat* OR assessment* OR assignment* OR test* OR exam* OR math* OR read* OR writ* OR spell*)

PsycInfo

TI,AB (self-efficacy OR "self efficacy") AND TI,AB (reciproc* OR longitud* OR "reverse caus*" OR circular OR lag* OR panel OR "growth curve" OR bidirect* OR bi-direct*) AND TI,AB (academi* OR universit* OR school* OR educat* OR assessment* OR assignment* OR test* OR exam* OR math* OR read* OR writ* OR spell*)

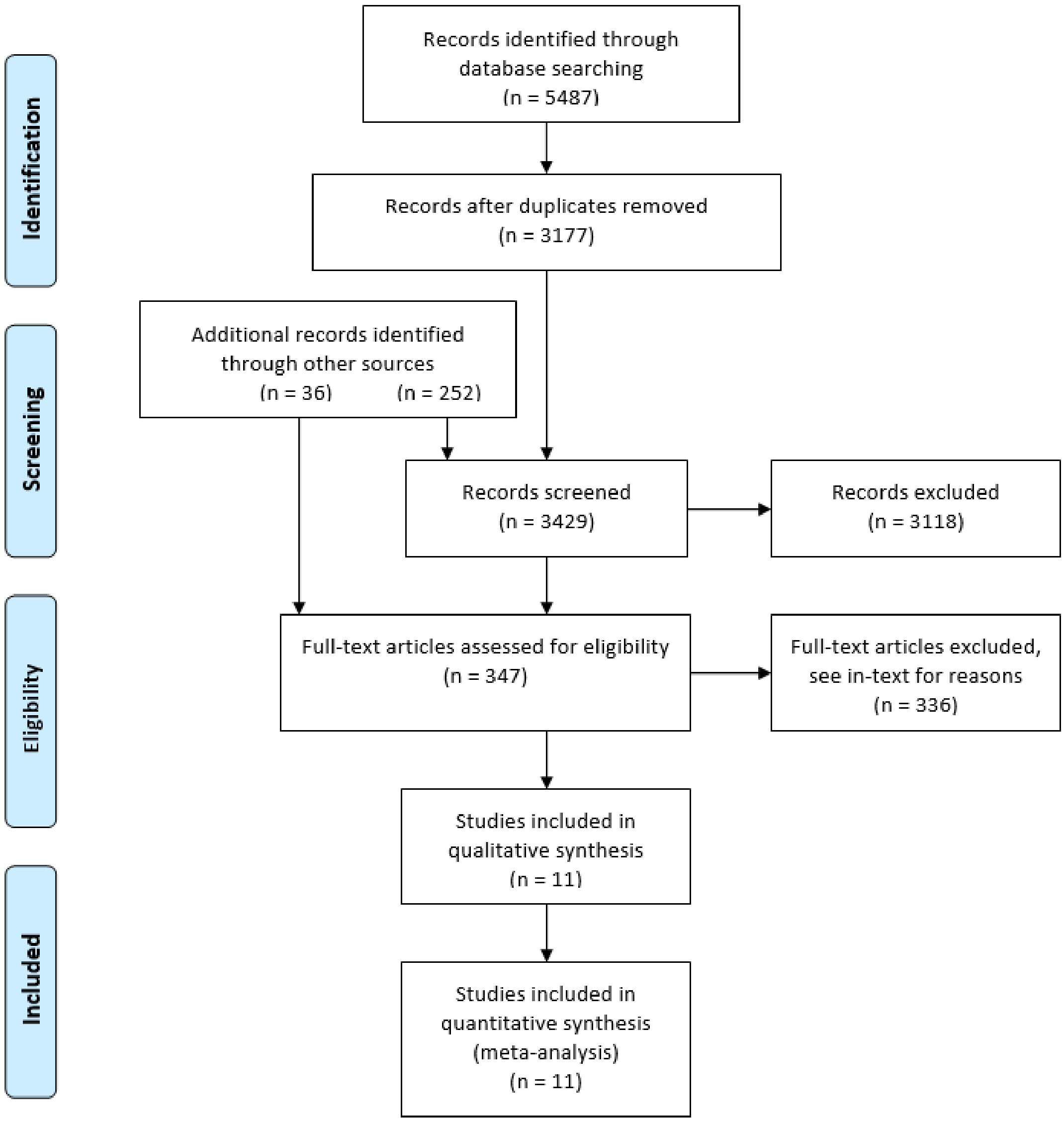
Web of Science

TS=((self-efficacy OR "self efficacy") AND (reciproc* OR longitud* OR "reverse caus*" OR circular OR lag* OR panel OR "growth curve" OR bidirect* OR bi-direct*) AND (academi* OR universit* OR school* OR educat* OR assessment* OR assignment* OR test* OR exam* OR math* OR read* OR writ* OR spell*))

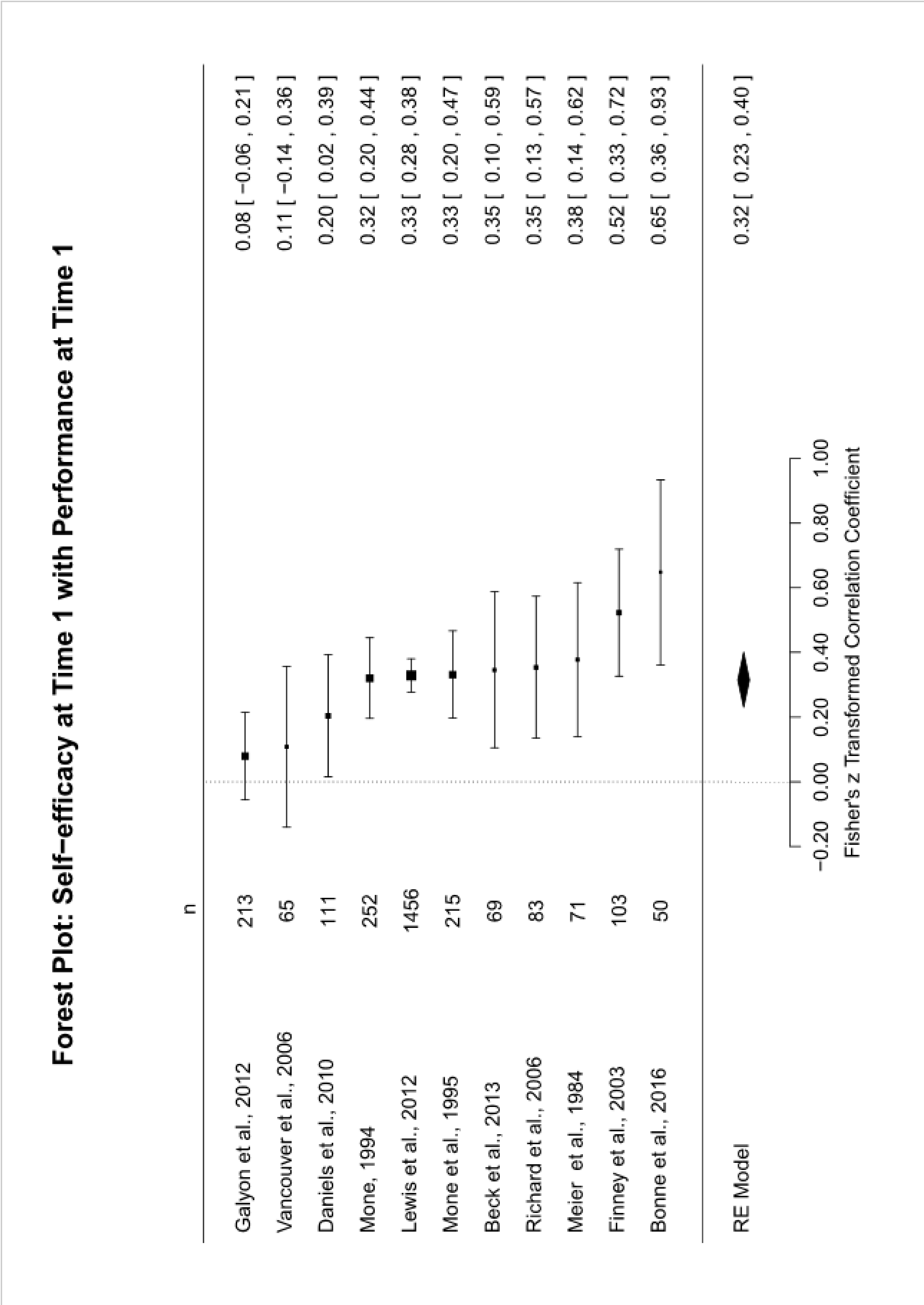
ERIC (via ProQuest)

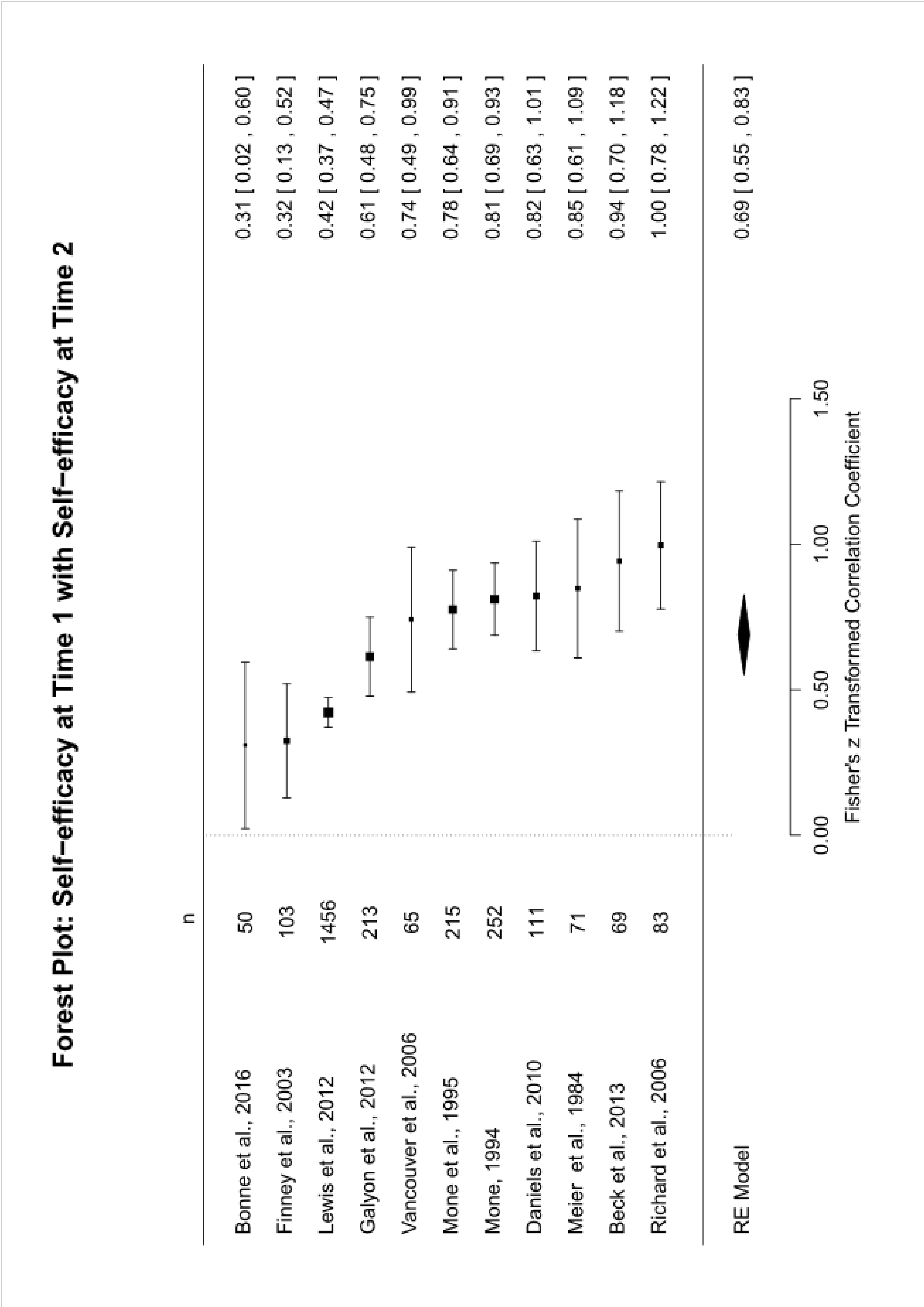
TI,AB (self-efficacy OR "self efficacy") AND TI,AB (reciproc* OR longitud* OR "reverse caus*" OR circular OR lag* OR panel OR "growth curve" OR bidirect* OR bi-direct*) AND TI,AB (academi* OR universit* OR school* OR educat* OR assessment* OR assignment* OR test* OR exam* OR math* OR read* OR writ* OR spell*)

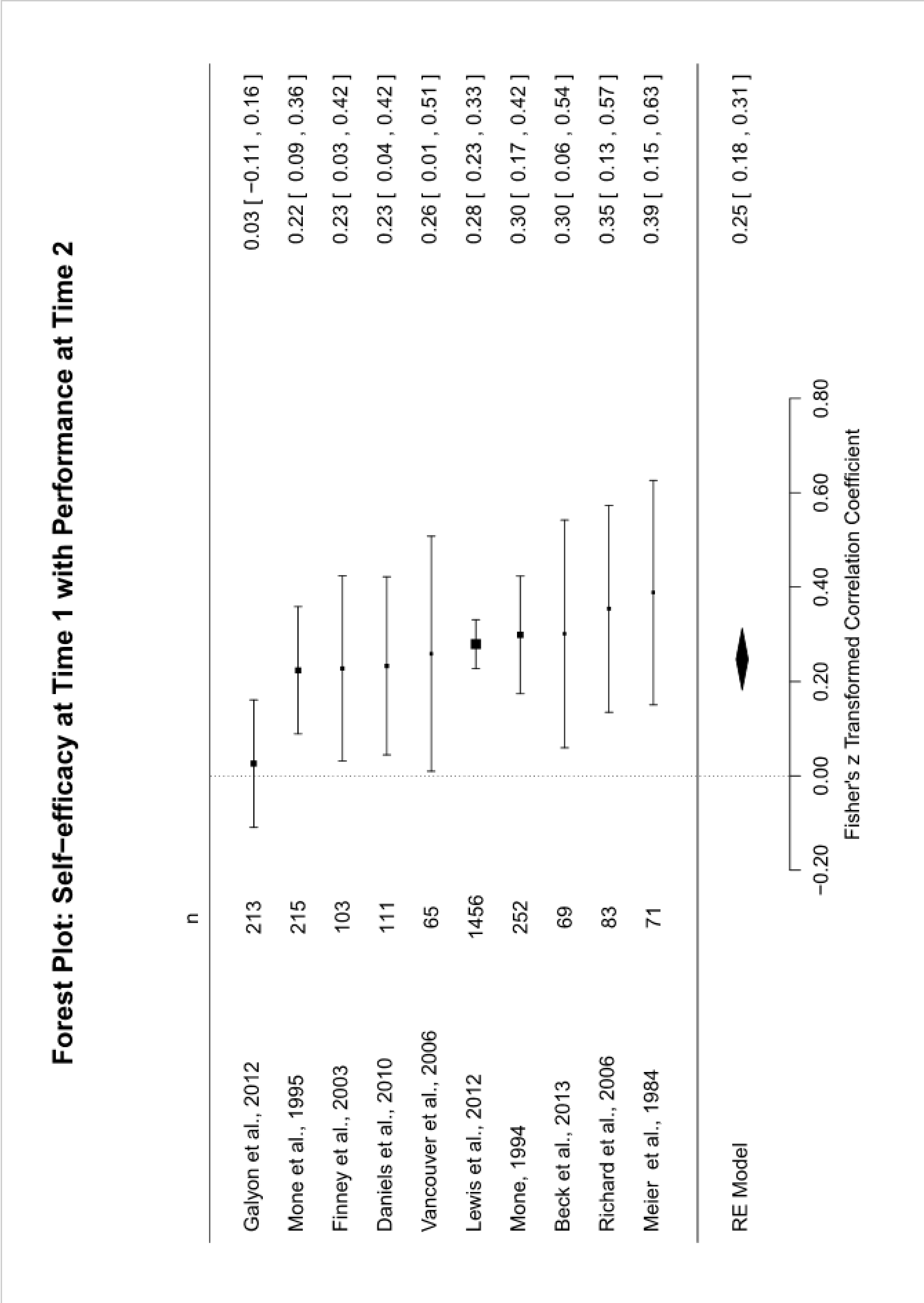
Appendix 2.2 Literature search flow chart



Appendix 2.3 Forest plots: Overall analysis

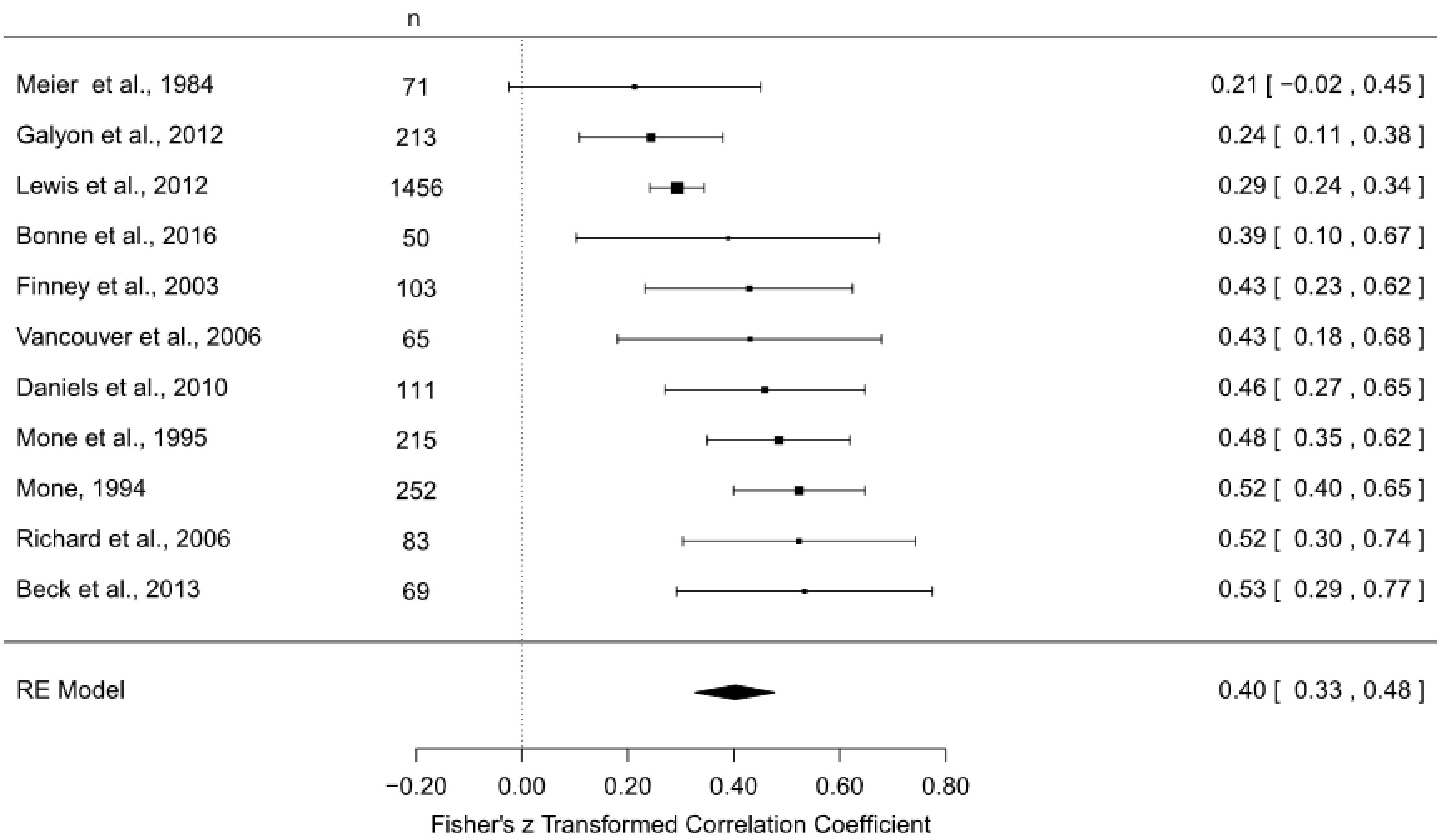

Forest Plot: Self–efficacy at Time 1 with Performance at Time 1
n
Galyon et al., 2012 213 0.08 [−0.06 , 0.21]
Vancouver et al., 2006 65 0.11 [−0.14 , 0.36]
Daniels et al., 2010 111 0.20 [0.02 , 0.39]
Mone, 1994 252 0.32 [0.20 , 0.44]
Lewis et al., 2012 1456 0.33 [0.28 , 0.38]
Mone et al., 1995 215 0.33 [0.20 , 0.47]
Beck et al., 2013 69 0.35 [0.10 , 0.59]
Richard et al., 2006 83 0.35 [0.13 , 0.57]
Meier et al., 1984 71 0.38 [0.14 , 0.62]
Finney et al., 2003 103 0.52 [0.33 , 0.72]
Bonne et al., 2016 50 0.65 [0.36 , 0.93]
RE Model 0.32 [0.23 , 0.40]
−0.20 0.00 0.20 0.40 0.60 0.80 1.00
Fisher's z Transformed Correlation Coefficient

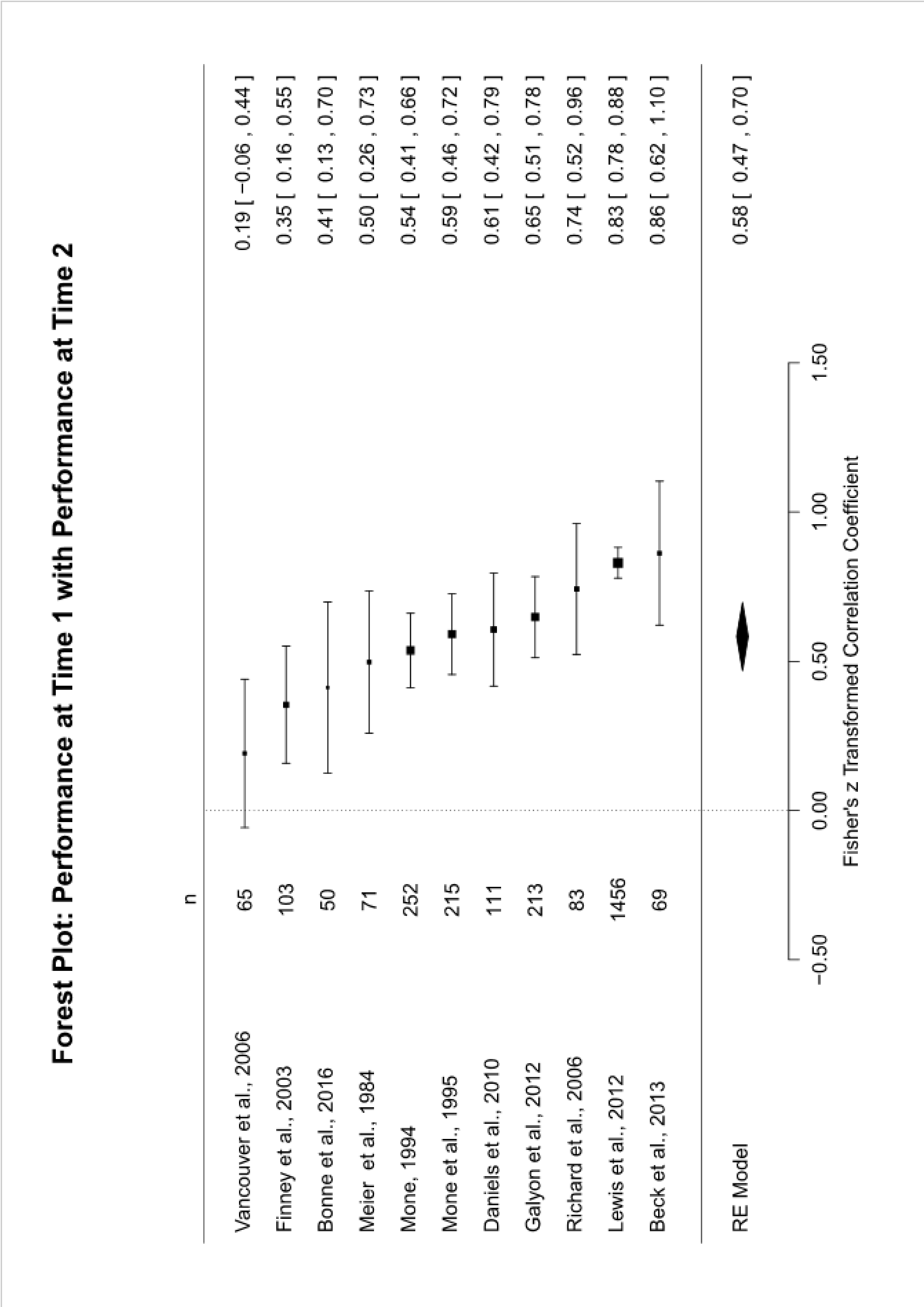

Forest Plot: Self–efficacy at Time 1 with Self–efficacy at Time 2
n
Bonne et al., 2016 50 0.31 [0.02 , 0.60]
Finney et al., 2003 103 0.32 [0.13 , 0.52]
Lewis et al., 2012 1456 0.42 [0.37 , 0.47]
Galyon et al., 2012 213 0.61 [0.48 , 0.75]
Vancouver et al., 2006 65 0.74 [0.49 , 0.99]
Mone et al., 1995 215 0.78 [0.64 , 0.91]
Mone, 1994 252 0.81 [0.69 , 0.93]
Daniels et al., 2010 111 0.82 [0.63 , 1.01]
Meier et al., 1984 71 0.85 [0.61 , 1.09]
Beck et al., 2013 69 0.94 [0.70 , 1.18]
Richard et al., 2006 83 1.00 [0.78 , 1.22]
RE Model 0.69 [0.55 , 0.83]
0.00
0.50
1.00
1.50
Fisher's z Transformed Correlation Coefficient

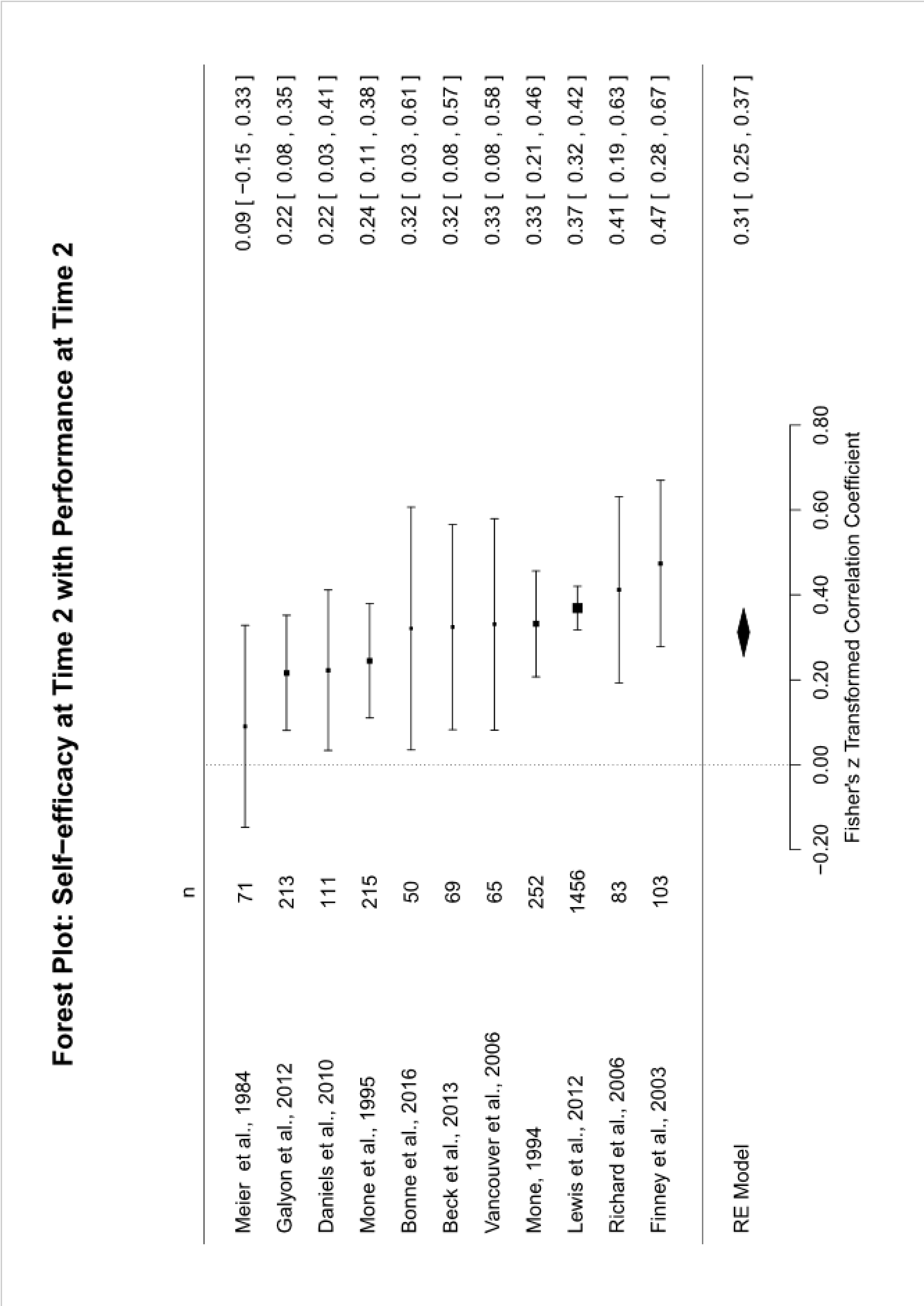


Forest Plot: Self−efficacy at Time 1 with Performance at Time 2
n
Galyon et al., 2012 213 0.03 [−0.11 , 0.16]
Mone et al., 1995 215 0.22 [0.09 , 0.36]
Finney et al., 2003 103 0.23 [0.03 , 0.42]
Daniels et al., 2010 111 0.23 [0.04 , 0.42]
Vancouver et al., 2006 65 0.26 [0.01 , 0.51]
Lewis et al., 2012 1456 0.28 [0.23 , 0.33]
Mone, 1994 252 0.30 [0.17 , 0.42]
Beck et al., 2013 69 0.30 [0.06 , 0.54]
Richard et al., 2006 83 0.35 [0.13 , 0.57]
Meier et al., 1984 71 0.39 [0.15 , 0.63]
RE Model 0.25 [0.18 , 0.31]
−0.20
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient

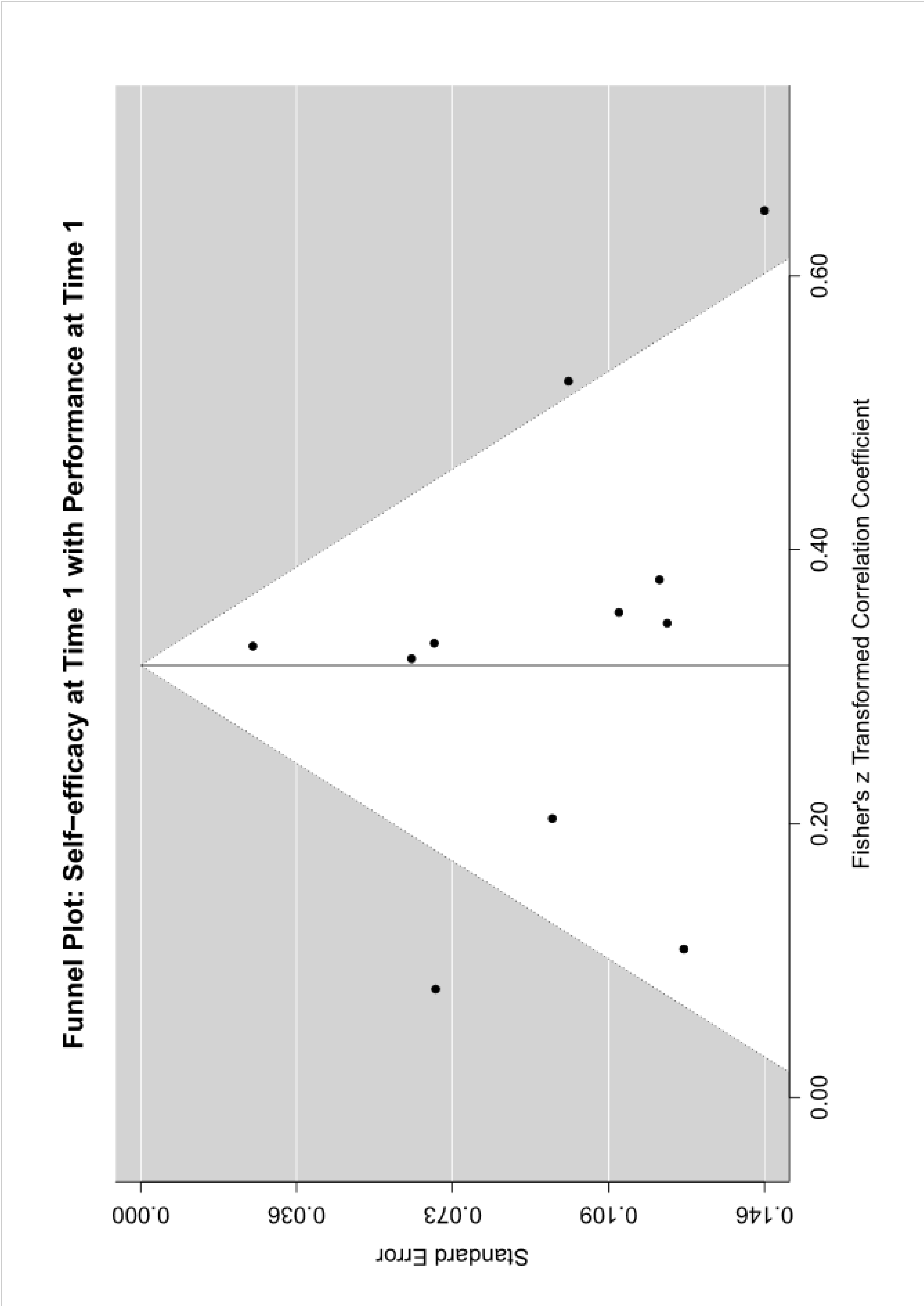
Forest Plot: Self−efficacy at Time 2 with Performance at Time 1

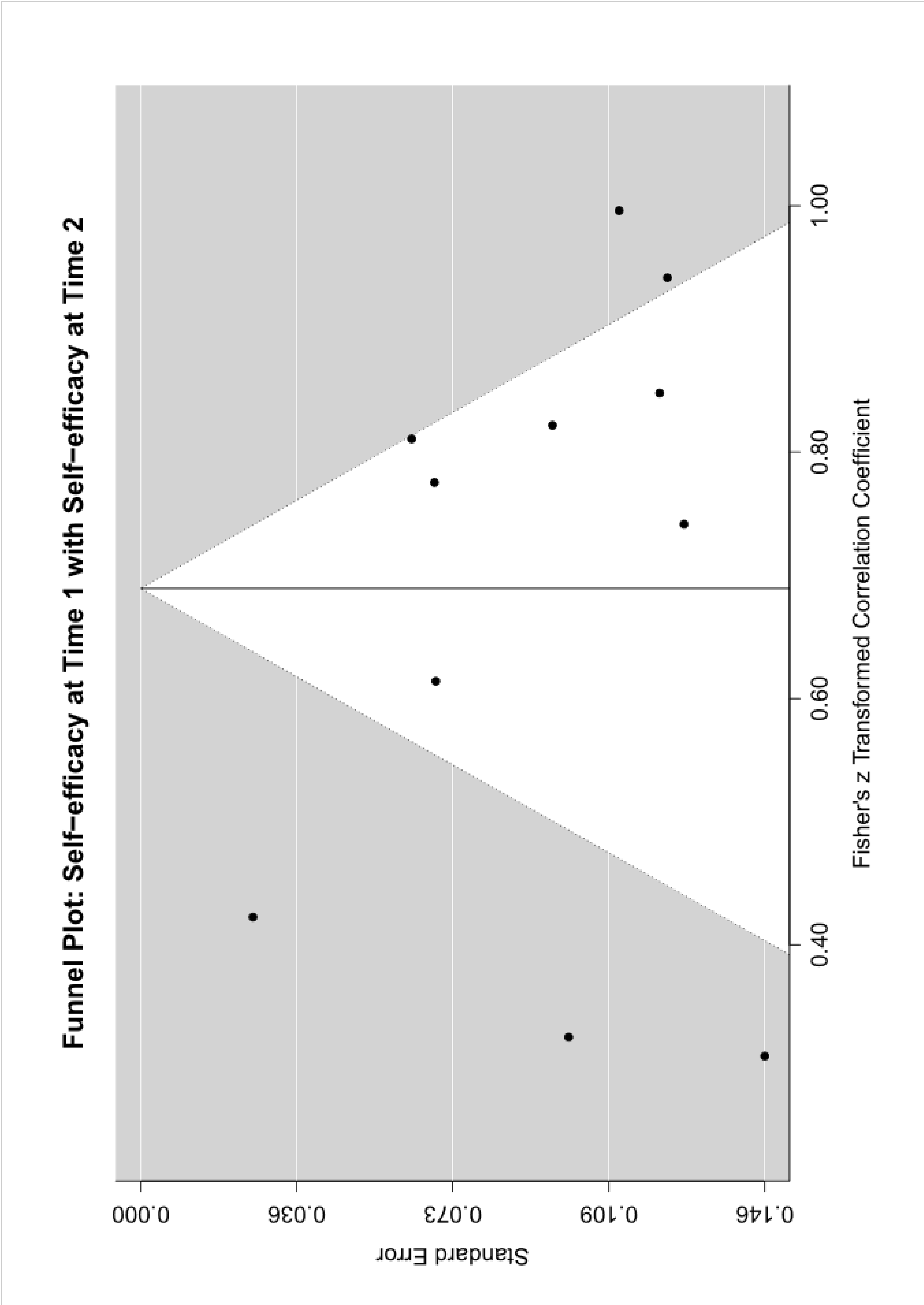


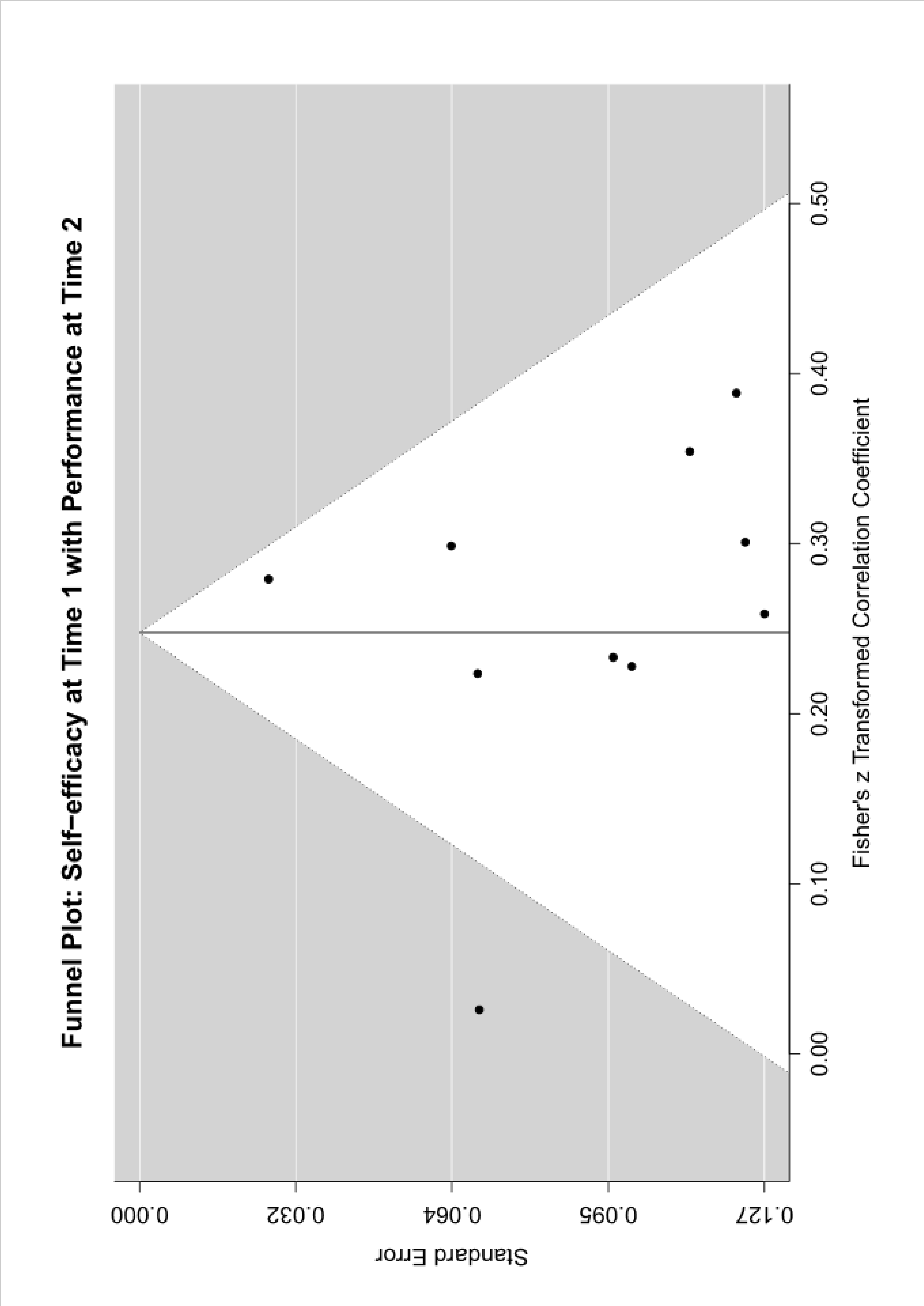

Forest Plot: Performance at Time 1 with Performance at Time 2
n
Vancouver et al., 2006 65 0.19 [−0.06 , 0.44]
Finney et al., 2003 103 0.35 [0.16 , 0.55]
Bonne et al., 2016 50 0.41 [0.13 , 0.70]
Meier et al., 1984 71 0.50 [0.26 , 0.73]
Mone, 1994 252 0.54 [0.41 , 0.66]
Mone et al., 1995 215 0.59 [0.46 , 0.72]
Daniels et al., 2010 111 0.61 [0.42 , 0.79]
Galyon et al., 2012 213 0.65 [0.51 , 0.78]
Richard et al., 2006 83 0.74 [0.52 , 0.96]
Lewis et al., 2012 1456 0.83 [0.78 , 0.88]
Beck et al., 2013 69 0.86 [0.62 , 1.10]
RE Model 0.58 [0.47 , 0.70]
−0.50
0.00
0.50
1.00
1.50
Fisher's z Transformed Correlation Coefficient


Forest Plot: Self−efficacy at Time 2 with Performance at Time 2
n
Meier et al., 1984 71 0.09 [−0.15 , 0.33]
Galyon et al., 2012 213 0.22 [0.08 , 0.35]
Daniels et al., 2010 111 0.22 [0.03 , 0.41]
Mone et al., 1995 215 0.24 [0.11 , 0.38]
Bonne et al., 2016 50 0.32 [0.03 , 0.61]
Beck et al., 2013 69 0.32 [0.08 , 0.57]
Vancouver et al., 2006 65 0.33 [0.08 , 0.58]
Mone, 1994 252 0.33 [0.21 , 0.46]
Lewis et al., 2012 1456 0.37 [0.32 , 0.42]
Richard et al., 2006 83 0.41 [0.19 , 0.63]
Finney et al., 2003 103 0.47 [0.28 , 0.67]
RE Model 0.31 [0.25 , 0.37]
−0.20 0.00 0.20 0.40 0.60 0.80
Fisher's z Transformed Correlation Coefficient

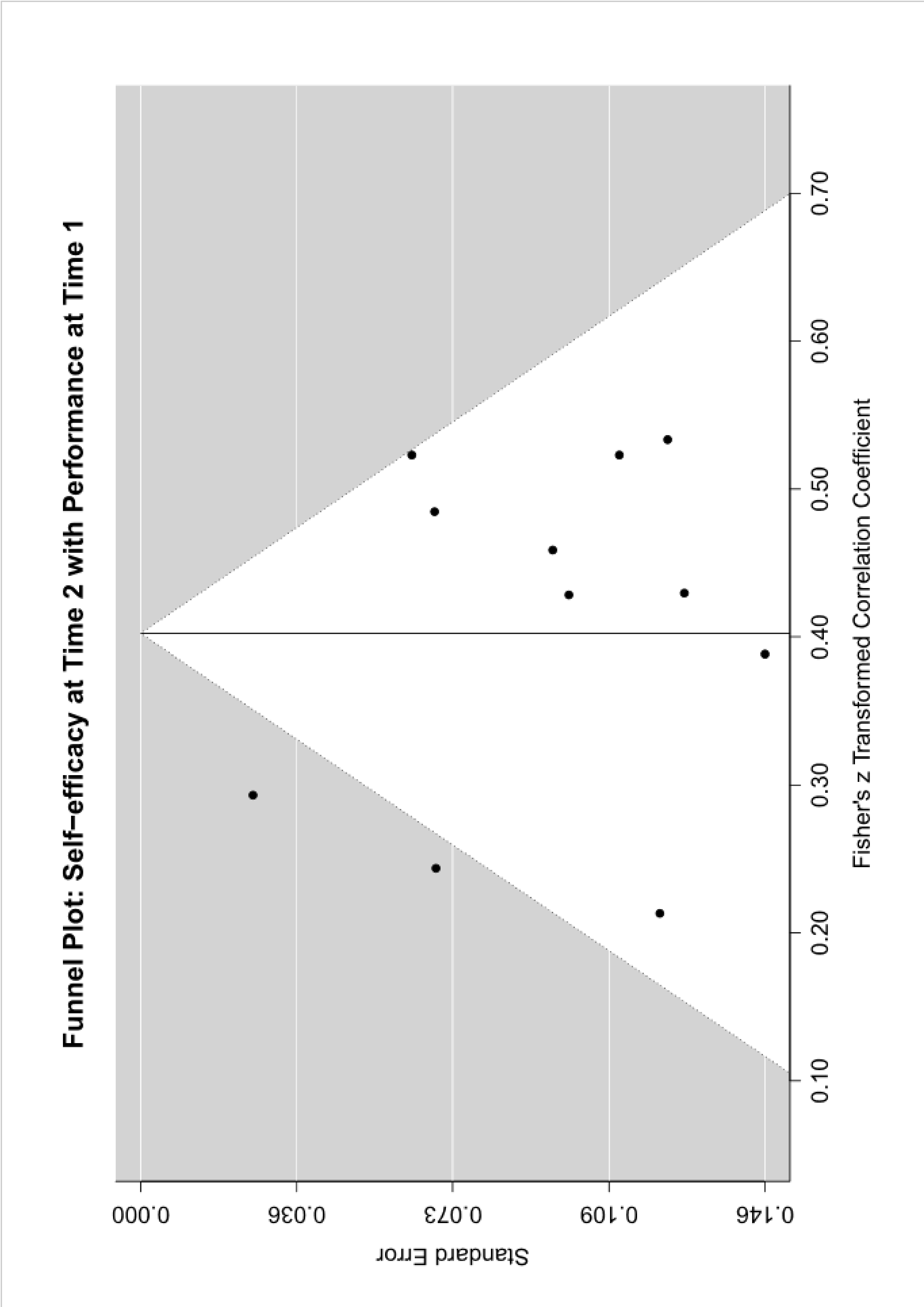
Appendix 2.4 Funnel plots: Overall analysis

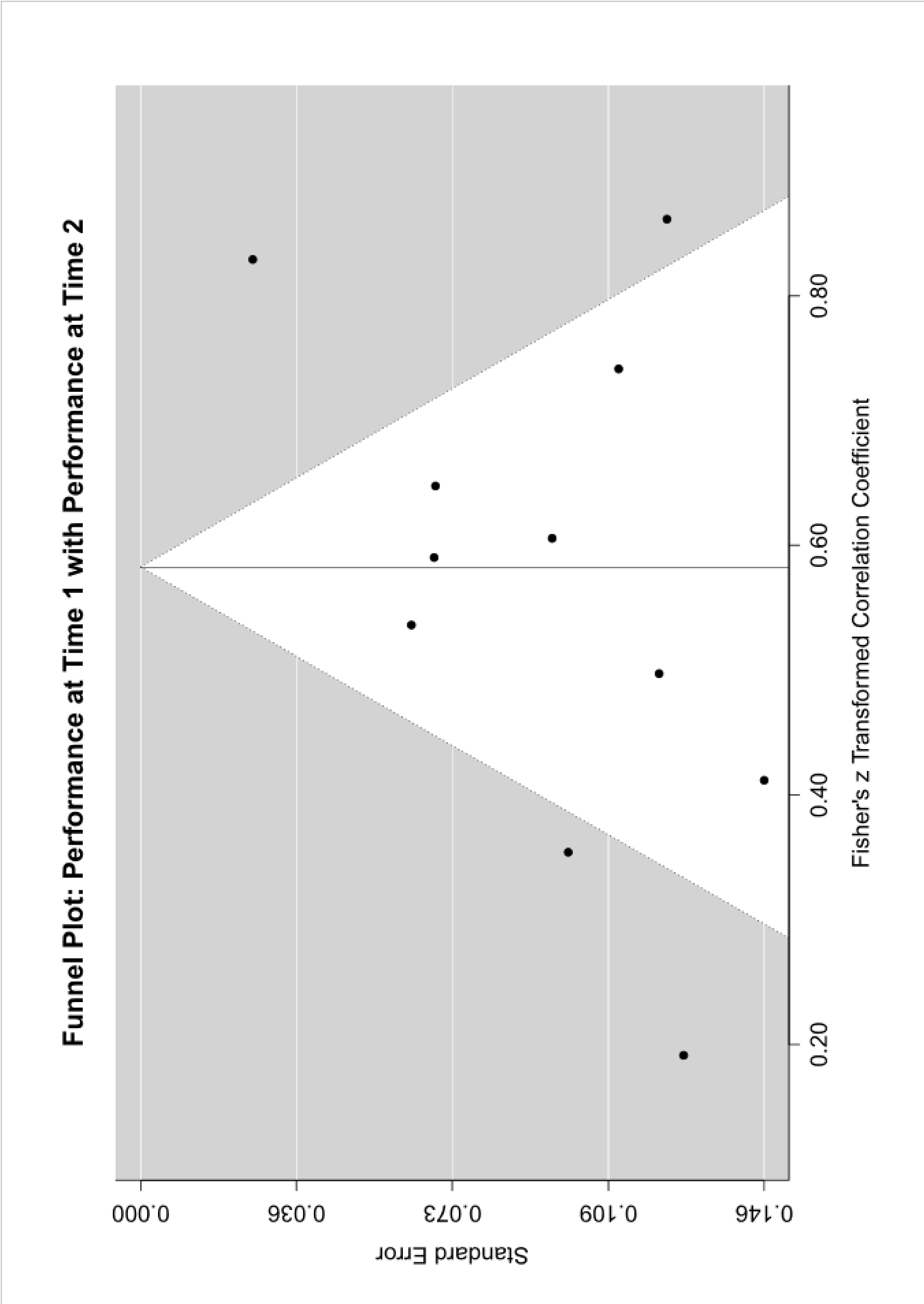

Funnel Plot: Self−efficacy at Time 1 with Performance at Time 1
Standard Error
0.000
0.036
0.073
0.109
0.146
0.00
0.20
0.40
0.60
Fisher's z Transformed Correlation Coefficient

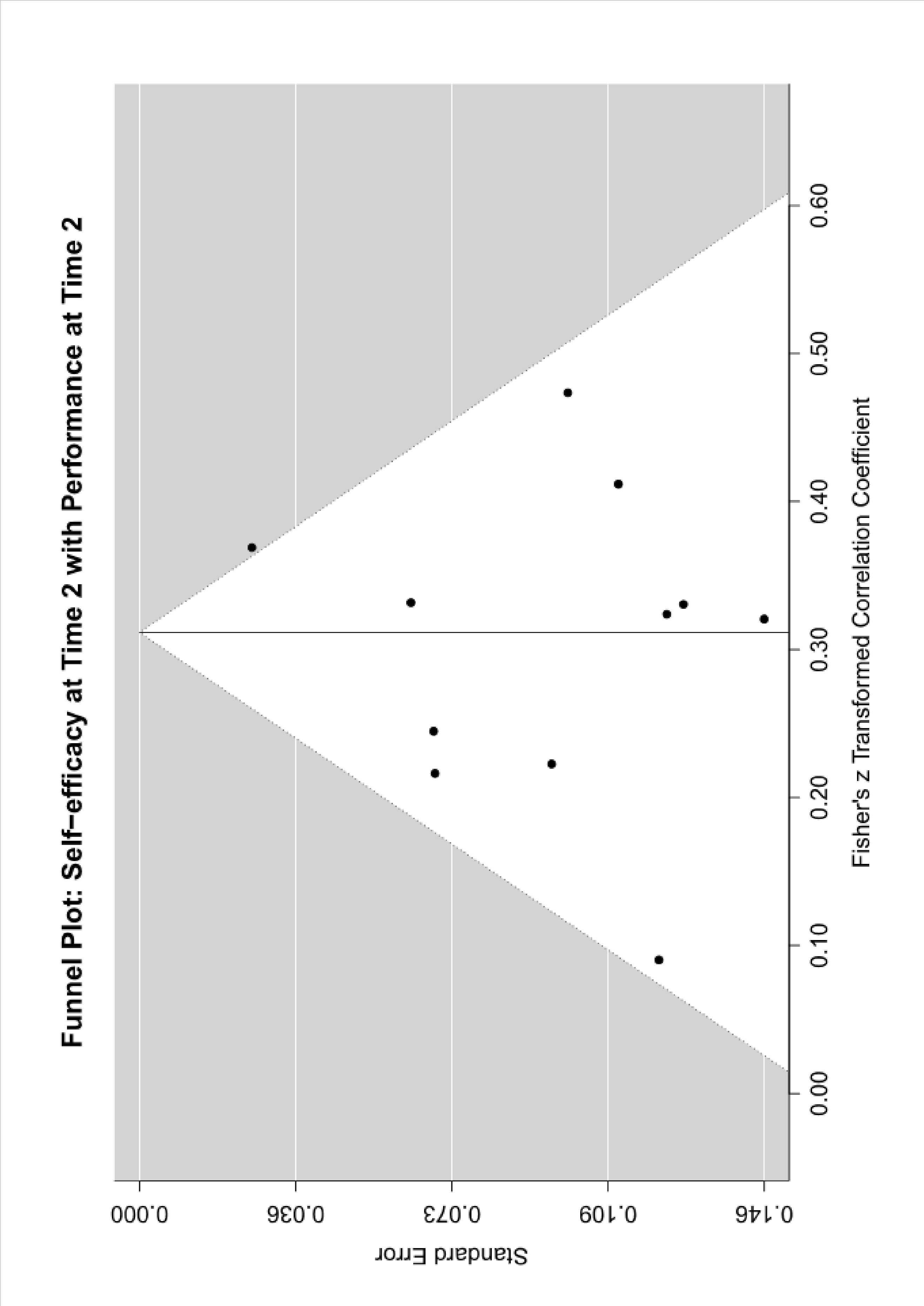

Funnel Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.40
0.60
0.80
1.00



Funnel Plot: Self−efficacy at Time 1 with Performance at Time 2
Standard Error
0.000
0.032
0.064
0.095
0.127
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50


Funnel Plot: Self−efficacy at Time 2 with Performance at Time 1
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.10
0.20
0.30
0.40
0.50
0.60
0.70


Funnel Plot: Performance at Time 1 with Performance at Time 2
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.20
0.40
0.60
0.80


Funnel Plot: Self−efficacy at Time 2 with Performance at Time 2
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50
0.60

Appendix 2.5 Meta-regression analyses for moderator variables

		SE1P1	SE1P2	SE2P1	SE2P2	SE1SE2	P1P2
Age group	Q	1.37	0.18	1.92	3.30	7.03**	0.59
	R^2	0	0	25.26	87.09	45.38	0.98
Lag	Q	1.16	0.43	2.46	3.78	5.98*	1.08
	R^2	0	0	25.33	92.87	40.47	6.85
Specificity match	Q	0.02	.95	0.06	2.06	0.78	10.86**
	R^2	0	2.18	0	67.28	0	64.89
Scale	Q	0.01	0.68	4.26*	2.83	5.55*	0.21
	R^2	0	0	63.42	48.44	40.05	0
Proportion males	Q	0.56	1.65	0.04	0.43	0.06	0.29
	R^2	0	36.47	0	0	0	0
Cohort	Q	0.14	1.25	0.26	1.13	1.36	0.97
	R^2	0	0	2.22	8.83	5.25	3.80

Note: Q values, df=1; SE1P1 = self-efficacy at time 1 with performance at time 1; SE1P2 = self-efficacy at time 1 with performance at time 2; SE2P1 = self-efficacy at time 2 with performance at time 1; SE2P2 = self-efficacy at time 2 with performance at time 2; SE1SE2 = self-efficacy at time 1 with self-efficacy at time 2; P1P2 = performance at time 1 with performance at time 2. * $p<.05$ ** $p<.01$

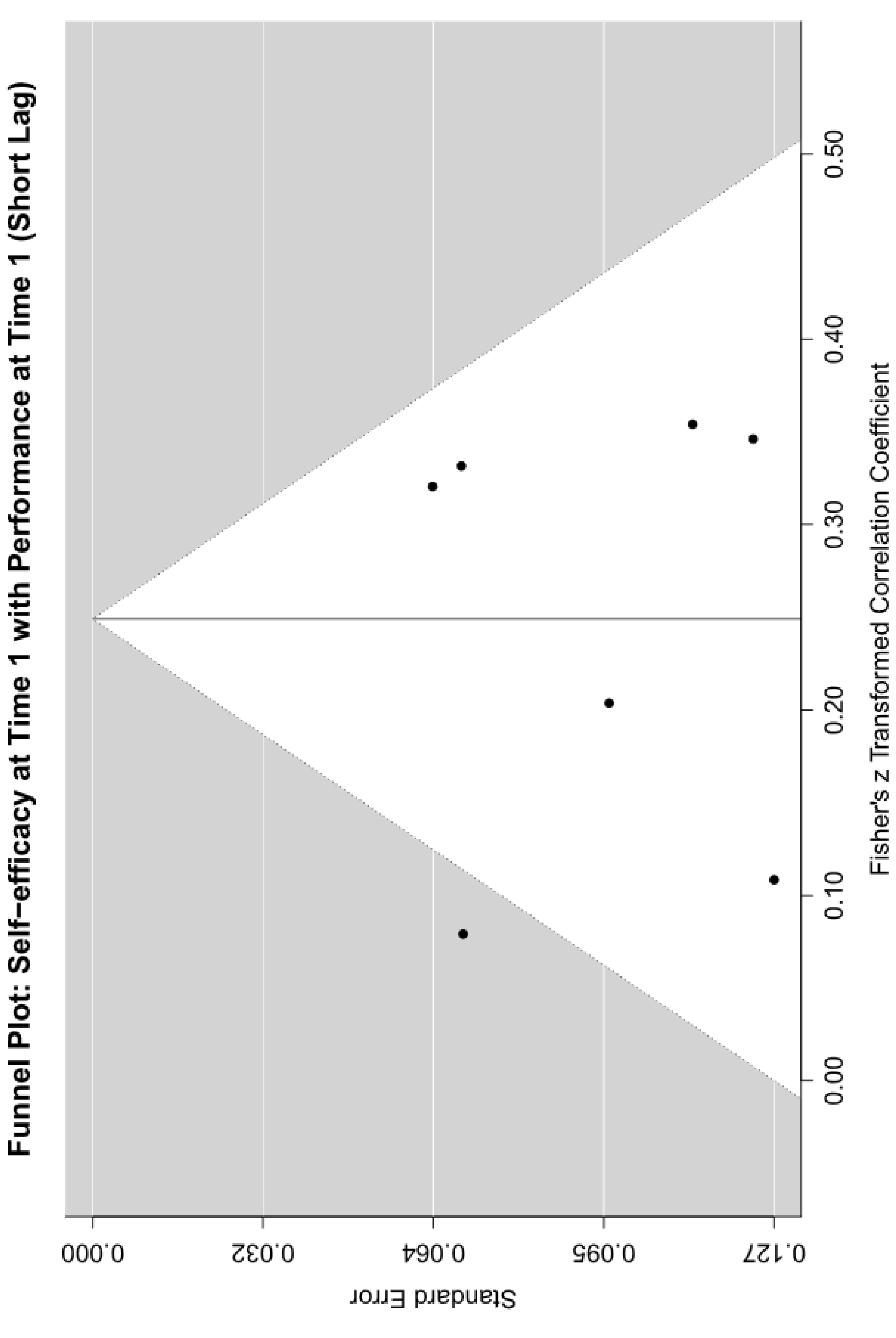
Appendix 2.6 Heterogeneity and power test values for moderator subgroups

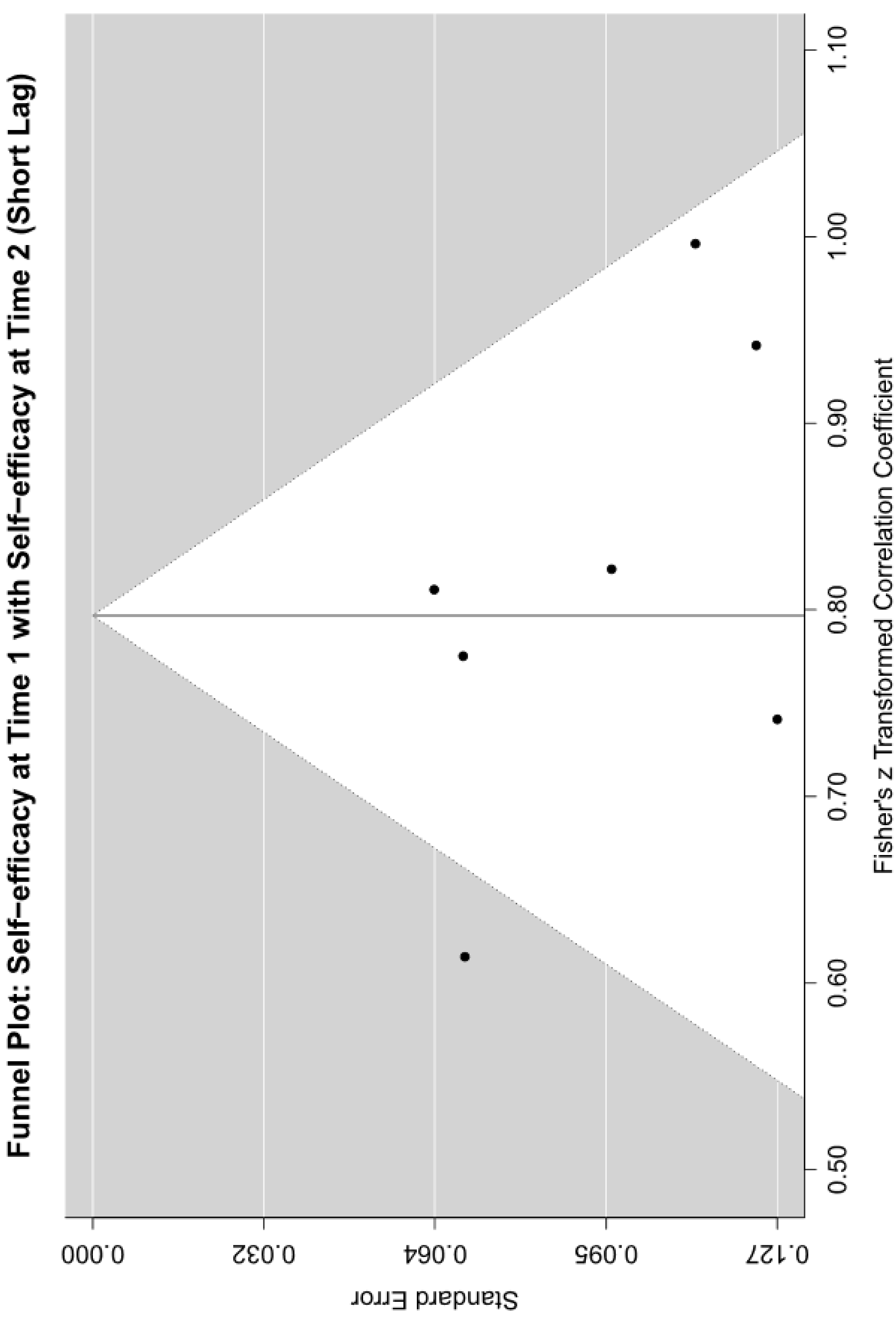
		n		SE1P1	SE1P2	SE2P1	SE2P2	SE1SE2	P1P2
Age group	Adults	1182	I^2	58.10	42.31	45.69	12.00	77.57	72.00
			Q	19.18*	13.239	14.75	9.93	31.47***	23.64**
			Power	96 [85,99]	70 [32,91]	99 [99,99]	93 [76,98]	99 [99,99]	99 [99,99]
	Children	1506	I^2	78.29		0	0	0	87.39
			Q	4.61*	0	0.41	.11	0.58	7.93**
			Power	99 [99,99]	99 [99,99]	99 [99,99]	99 [99,99]	99 [99,99]	99 [99,99]
Lag	Short	1008	I^2	48.79	48.53	48.30	0	49.85	73.99
			Q	11.73	11.43	11.39	3.80	11.76	17.97**
			Power	94 [72,99]	73 [31,93]	99 [99,99]	88 [64, 97]	99 [99,99]	99 [99,99]
	Long	1680	I^2	59.4	0	0	58.98	85.61	85.89
			Q	7.82	1.07	2.60	6.41	13.75*	33.41***
			Power	99 [99,99]	99 [99,99]	99 [99,99]	99 [99,99]	99 [99,99]	99 [99,99]
Matched specificity	Yes	969	I^2	77.23	55.69	52.85	21.91	82.18	62.18

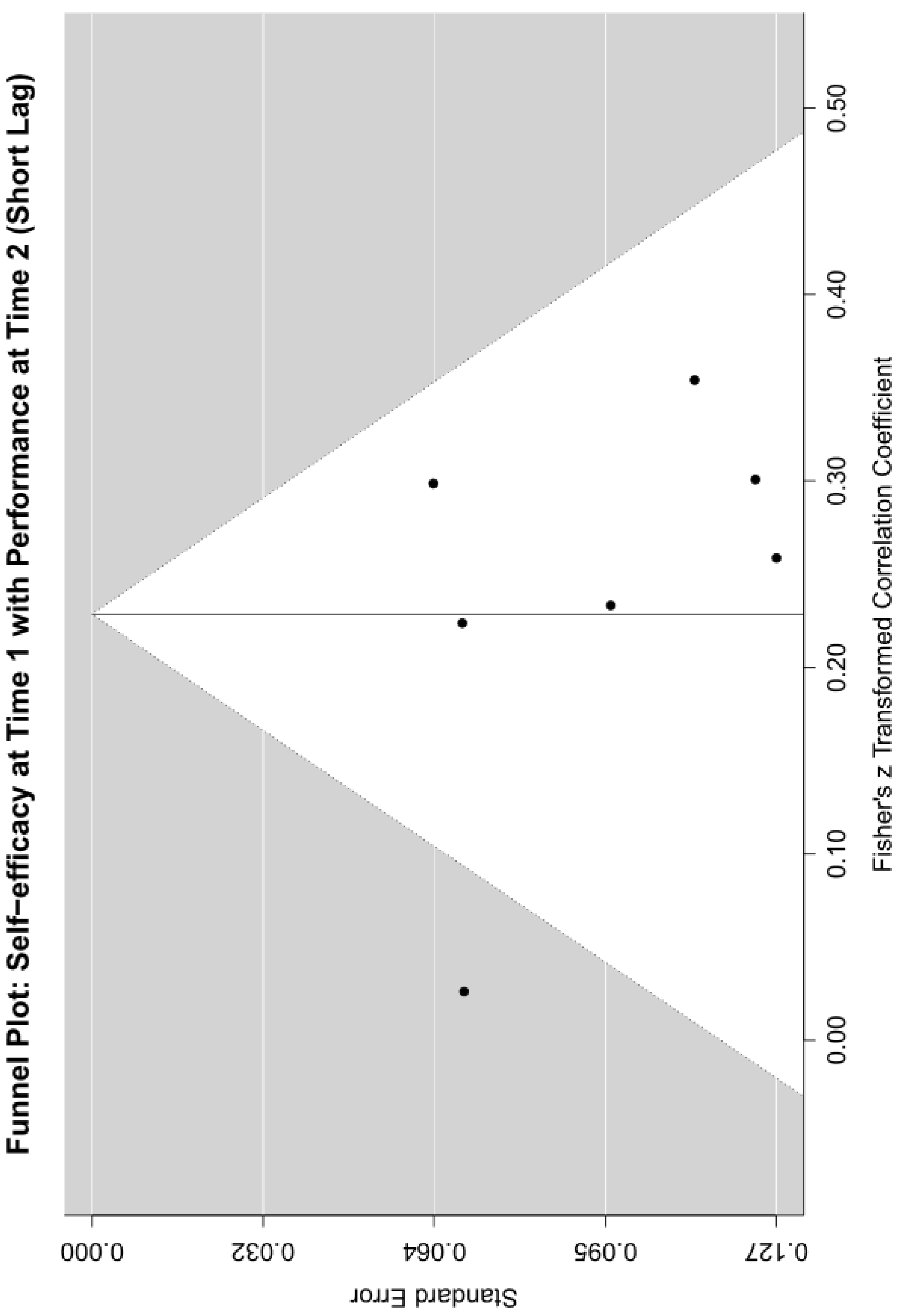
			Q	23.89***	11.43*	12.88*	8.19	28.30***	14.55*
			Power	99 [97,99]	76 [32, 95]	99 [99,99]	94 [76,99]	99 [99,99]	99 [99,99]
	No	1719	I^2	0	0	63.18	0	91.09	46.85
			Q	1.71	0.71	9.41*	2.45	52.44***	5.59
			Power	99 [92,99]	99 [93,99]	99 [99,99]	99 [96,99]	99 [99,99]	99 [99,99]
Scale	Unipolar	718	I^2	0	0	0.04	0	0	8.53
			Q	1.74	1.81	5.51	3.69	1.47	6.20
			Power	93 [71,99]	85 [51, 97]	99 [99,99]	91 [59, 99]	99 [99,99]	99 [99,99]
	Likert	1970	I^2	85.09	67.25	41.67	28.18	90.55	89.82
			Q	24.27***	12.68*	7.54	6.06	37.68***	51.71***
			Power	99 [99,99]	99 [82, 99]	99 [99,99]	99 [95,99]	99 [99,99]	99 [99,99]

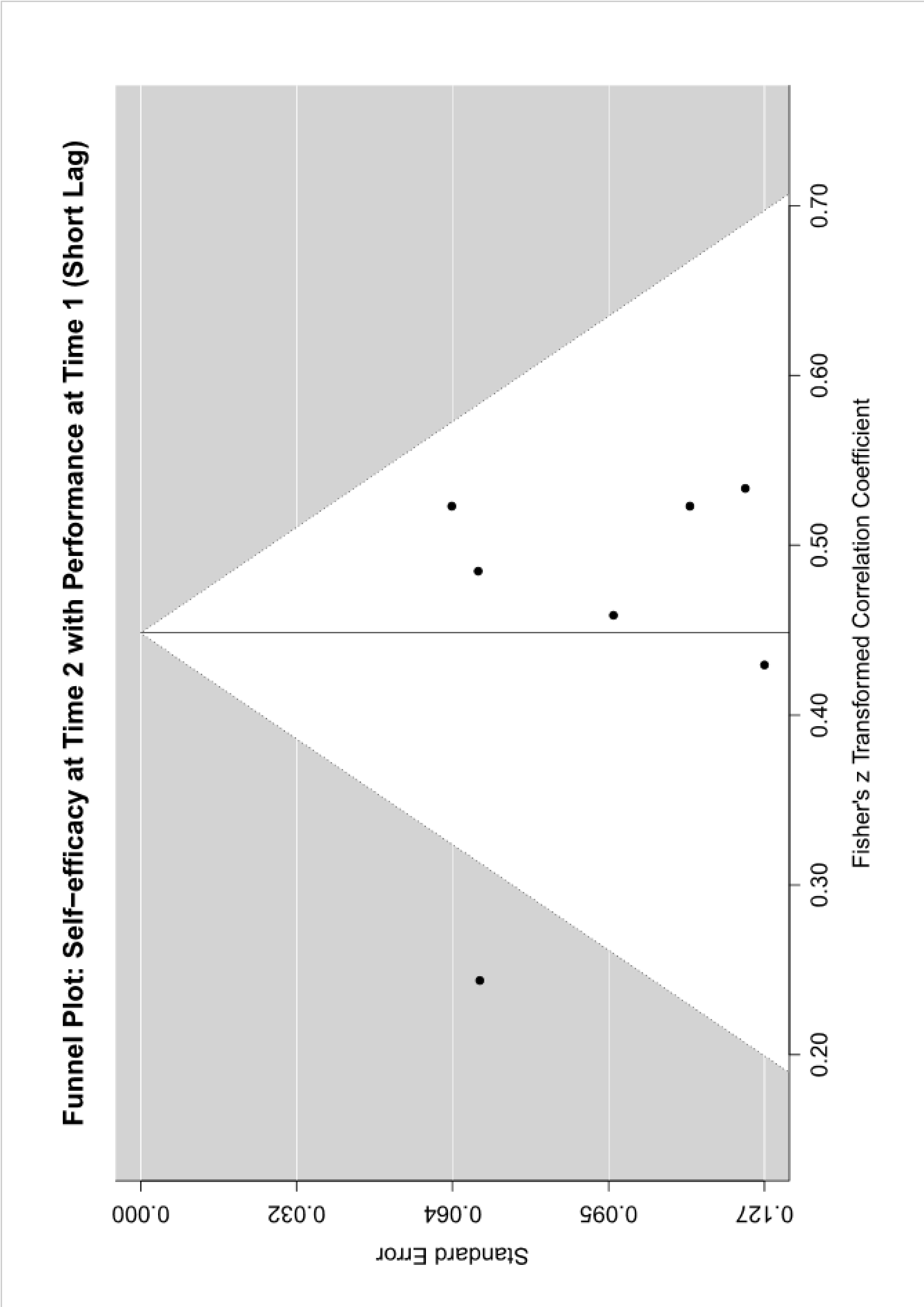
Note: r = pooled correlations after Fisher's Z transformations; SE= self-efficacy; SE1P1 = self-efficacy at time 1 with performance at time 1; SE1P2 = self-efficacy at time 1 with performance at time 2; SE2P1 = self-efficacy at time 2 with performance at time 1; SE2P2 = self-efficacy at time 2 with performance at time 2; SE1SE2 = self-efficacy at time 1 with self-efficacy at time 2; P1P2 = performance at time 1 with performance at time 2; * $p<.05$ ** $p<.01$ ***p<.001

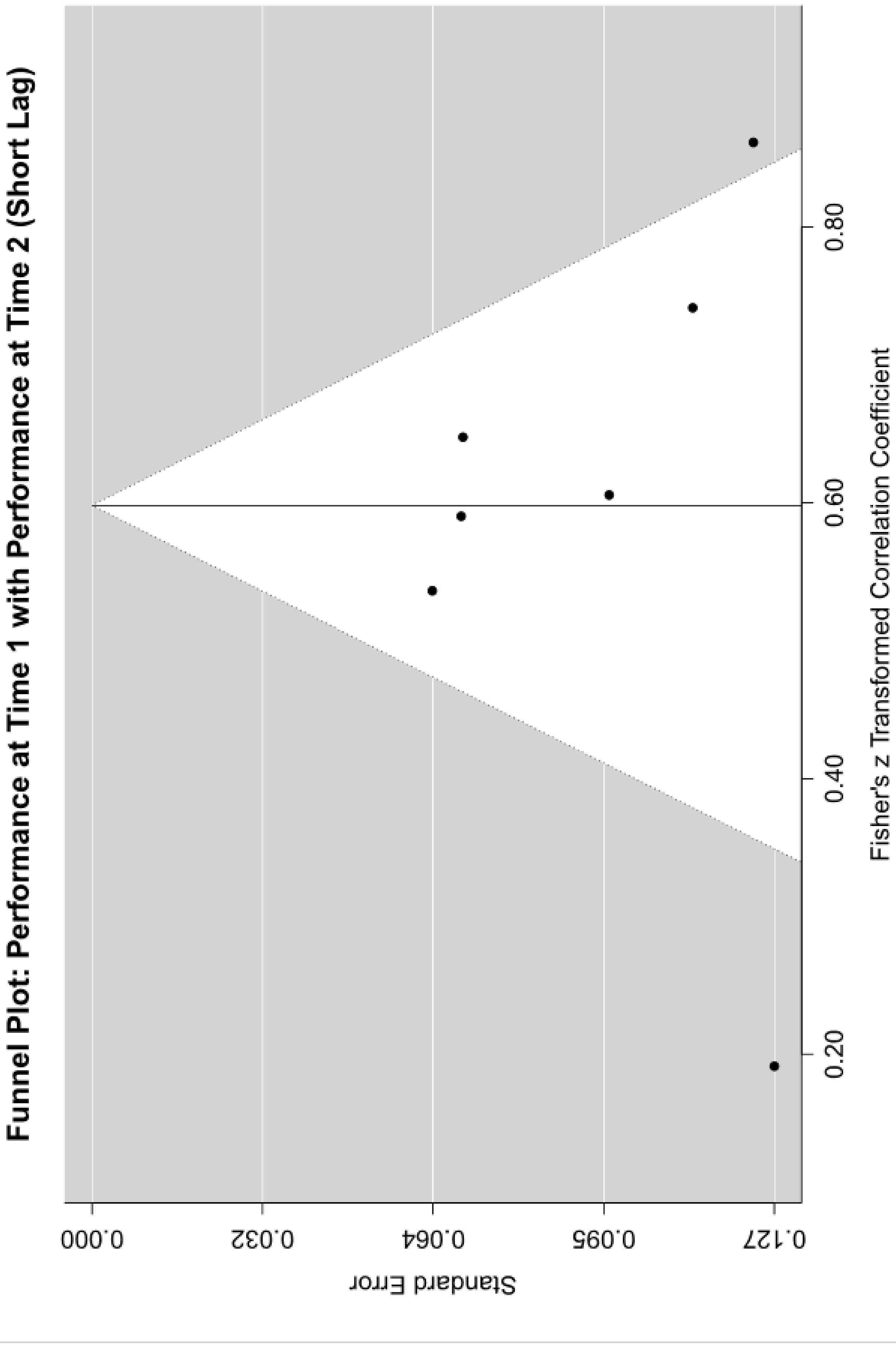
Appendix 2.7 Forest and funnel plots: moderator analyses

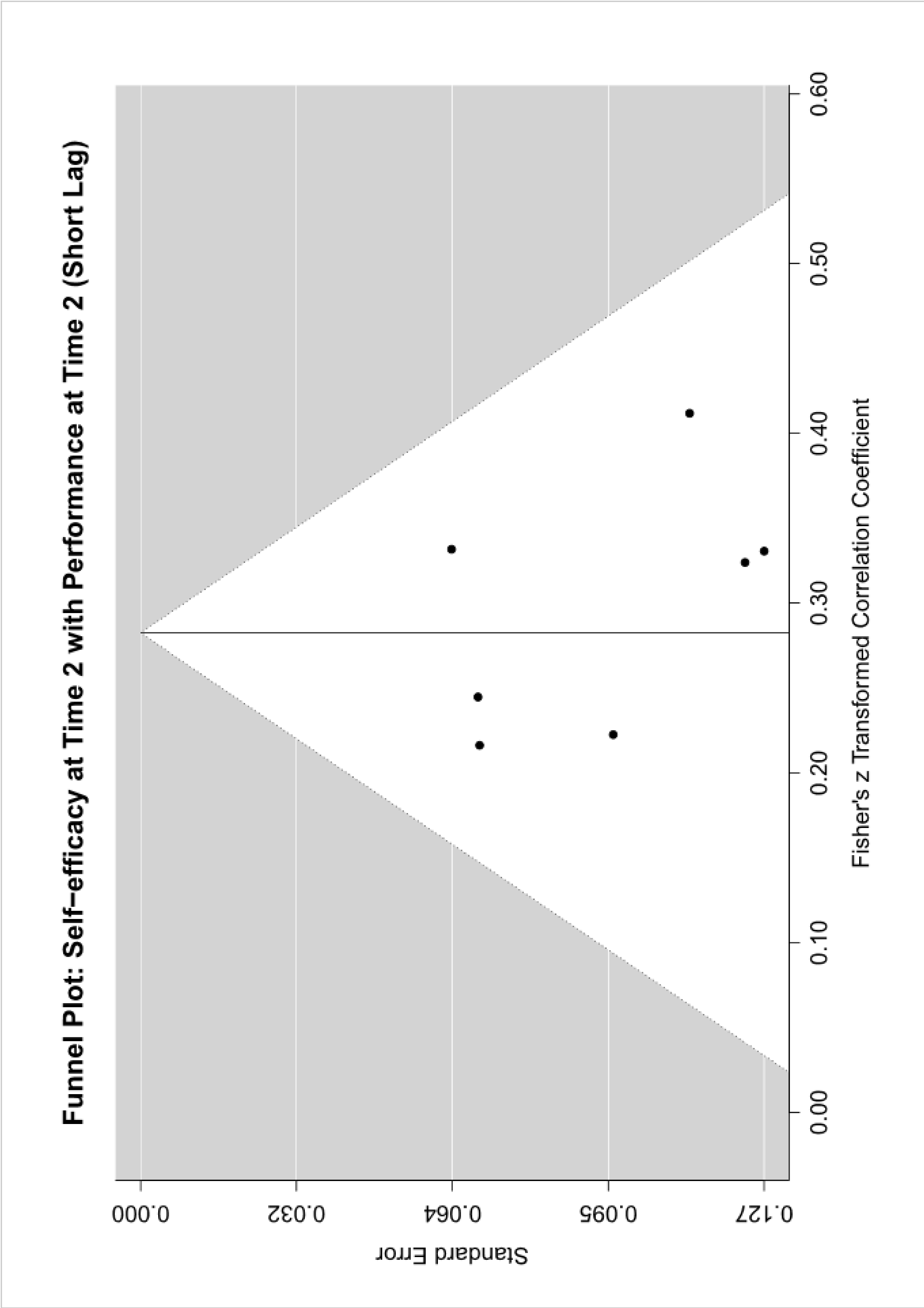

Funnel Plot: Self−efficacy at Time 1 with Performance at Time 1 (Short Lag)
Fisher's z Transformed Correlation Coefficient
Standard Error
0.000
0.032
0.064
0.095
0.127
0.00
0.10
0.20
0.30
0.40
0.50

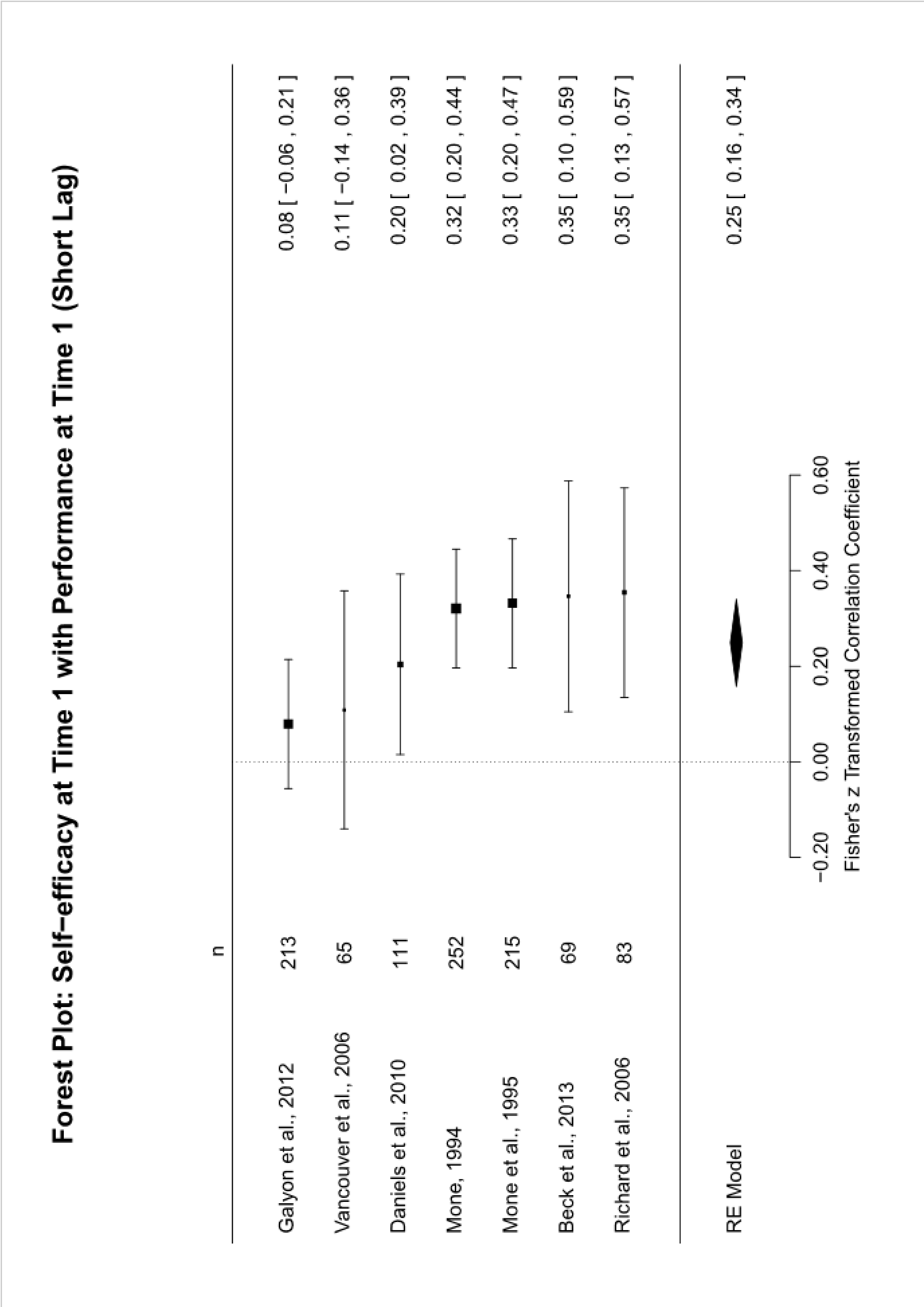

Funnel Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2 (Short Lag)
Standard Error
0.000
0.032
0.064
0.095
0.127
Fisher's z Transformed Correlation Coefficient
0.50
0.60
0.70
0.80
0.90
1.00
1.10

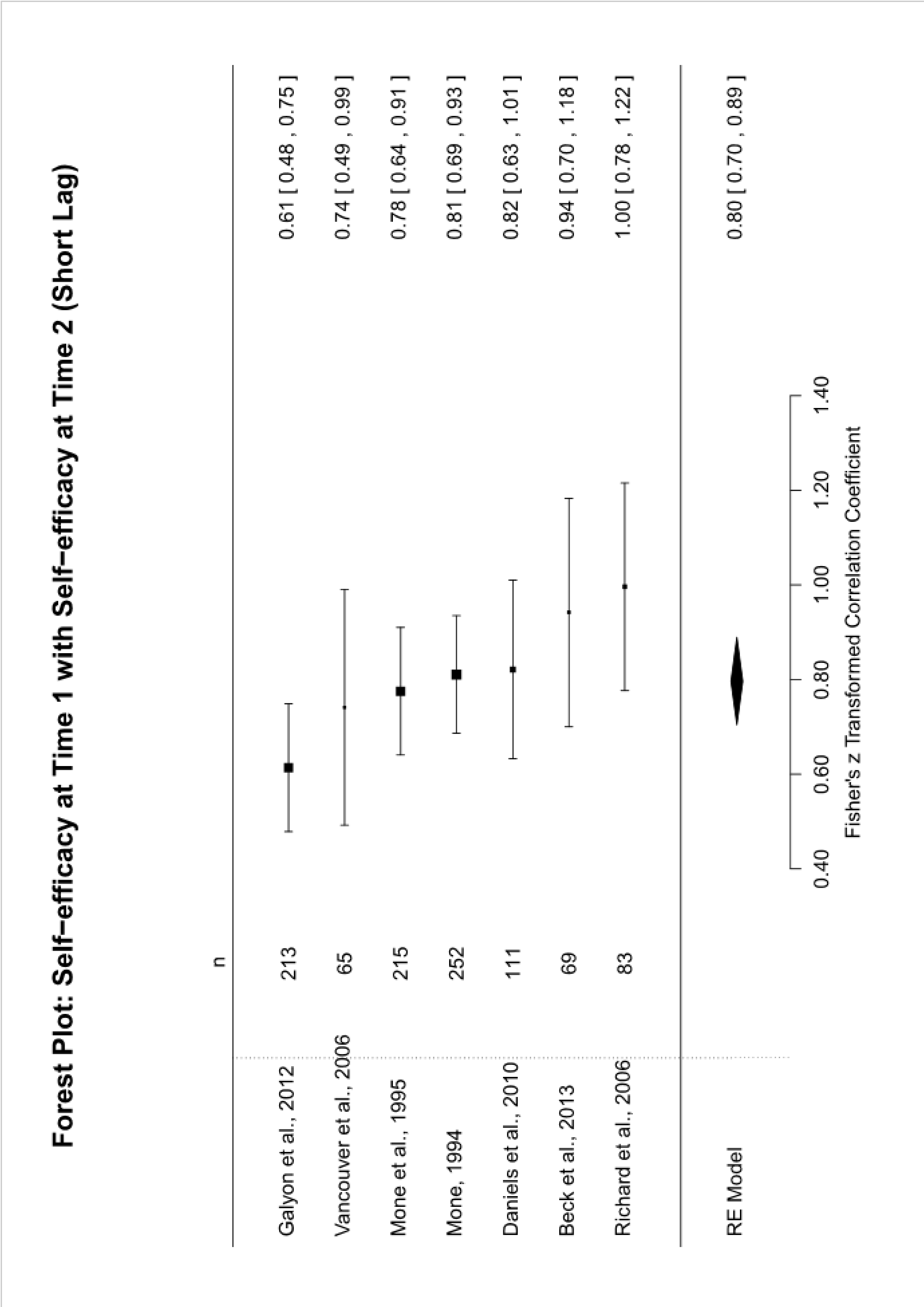

Funnel Plot: Self−efficacy at Time 1 with Performance at Time 2 (Short Lag)
Standard Error
0.000
0.032
0.064
0.095
0.127
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50


Funnel Plot: Self−efficacy at Time 2 with Performance at Time 1 (Short Lag)
Standard Error
0.000
0.032
0.064
0.095
0.127
Fisher's z Transformed Correlation Coefficient
0.20
0.30
0.40
0.50
0.60
0.70

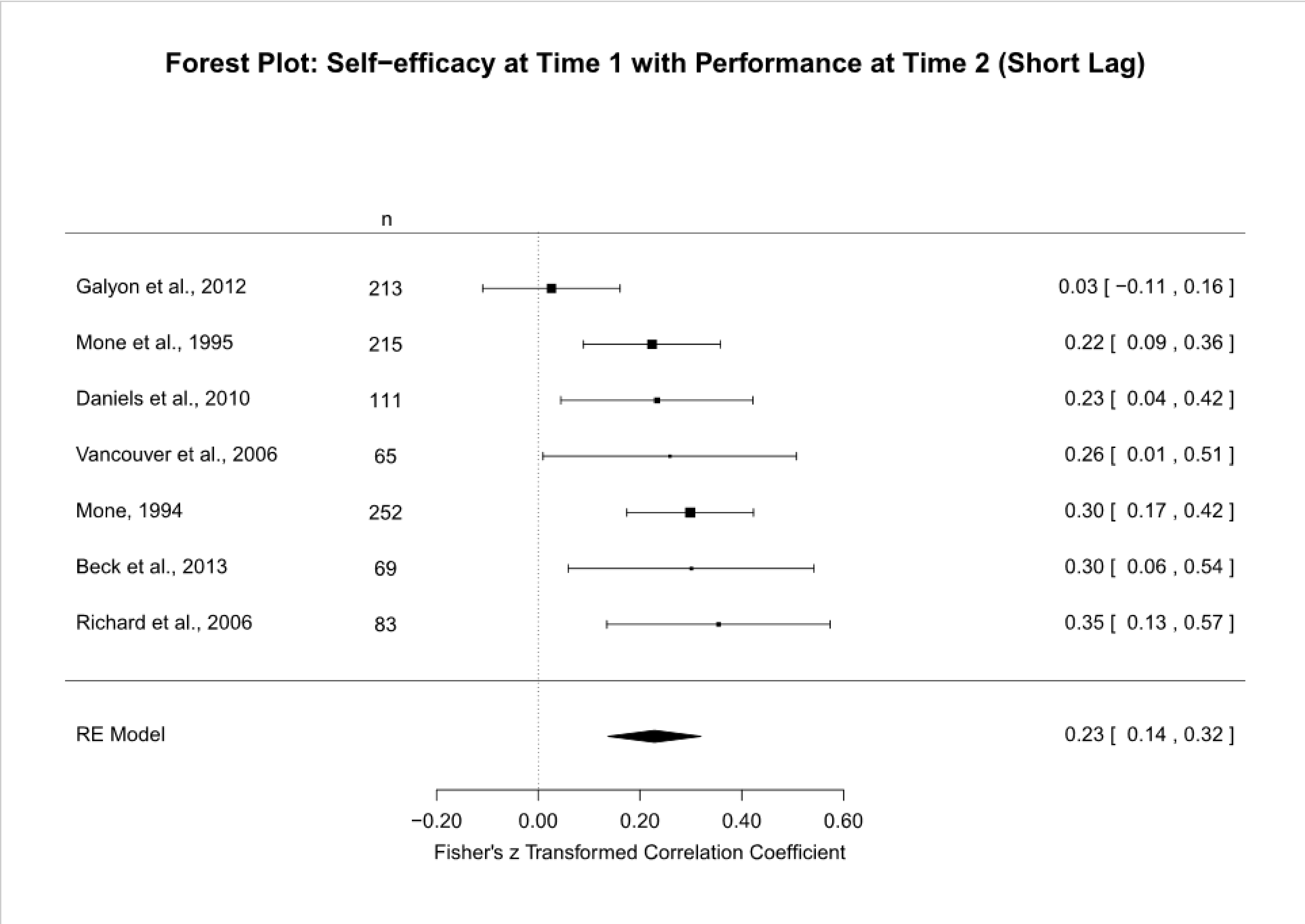

Funnel Plot: Performance at Time 1 with Performance at Time 2 (Short Lag)
Standard Error
0.000
0.032
0.064
0.095
0.127
Fisher's z Transformed Correlation Coefficient
0.20
0.40
0.60
0.80

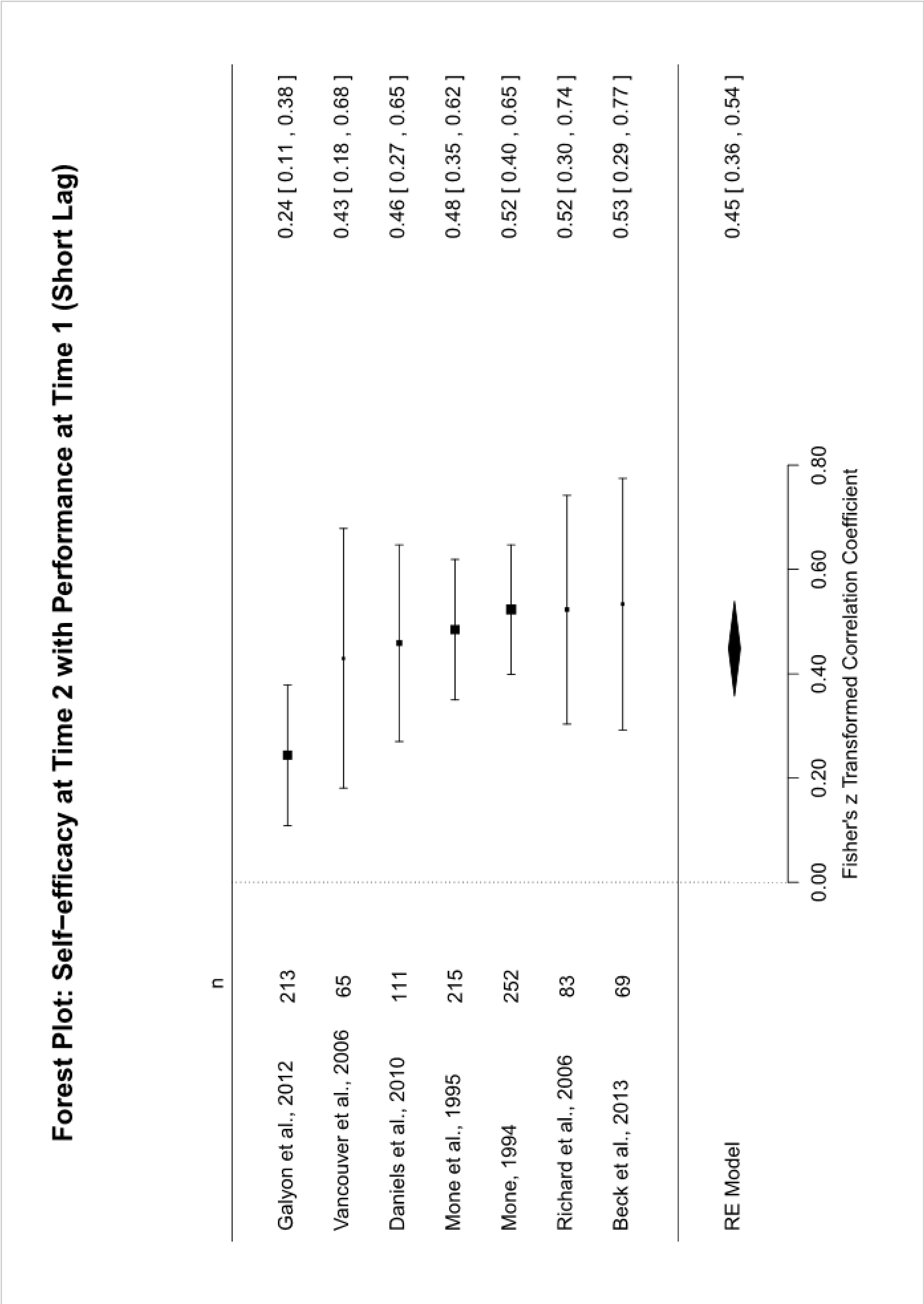

Funnel Plot: Self−efficacy at Time 2 with Performance at Time 2 (Short Lag)
Standard Error
0.000
0.032
0.064
0.095
0.127
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50
0.60


Forest Plot: Self–efficacy at Time 1 with Performance at Time 1 (Short Lag)
n
Galyon et al., 2012 213 0.08 [−0.06 , 0.21]
Vancouver et al., 2006 65 0.11 [−0.14 , 0.36]
Daniels et al., 2010 111 0.20 [0.02 , 0.39]
Mone, 1994 252 0.32 [0.20 , 0.44]
Mone et al., 1995 215 0.33 [0.20 , 0.47]
Beck et al., 2013 69 0.35 [0.10 , 0.59]
Richard et al., 2006 83 0.35 [0.13 , 0.57]
RE Model 0.25 [0.16 , 0.34]
−0.20
0.00
0.20
0.40
0.60
Fisher's z Transformed Correlation Coefficient

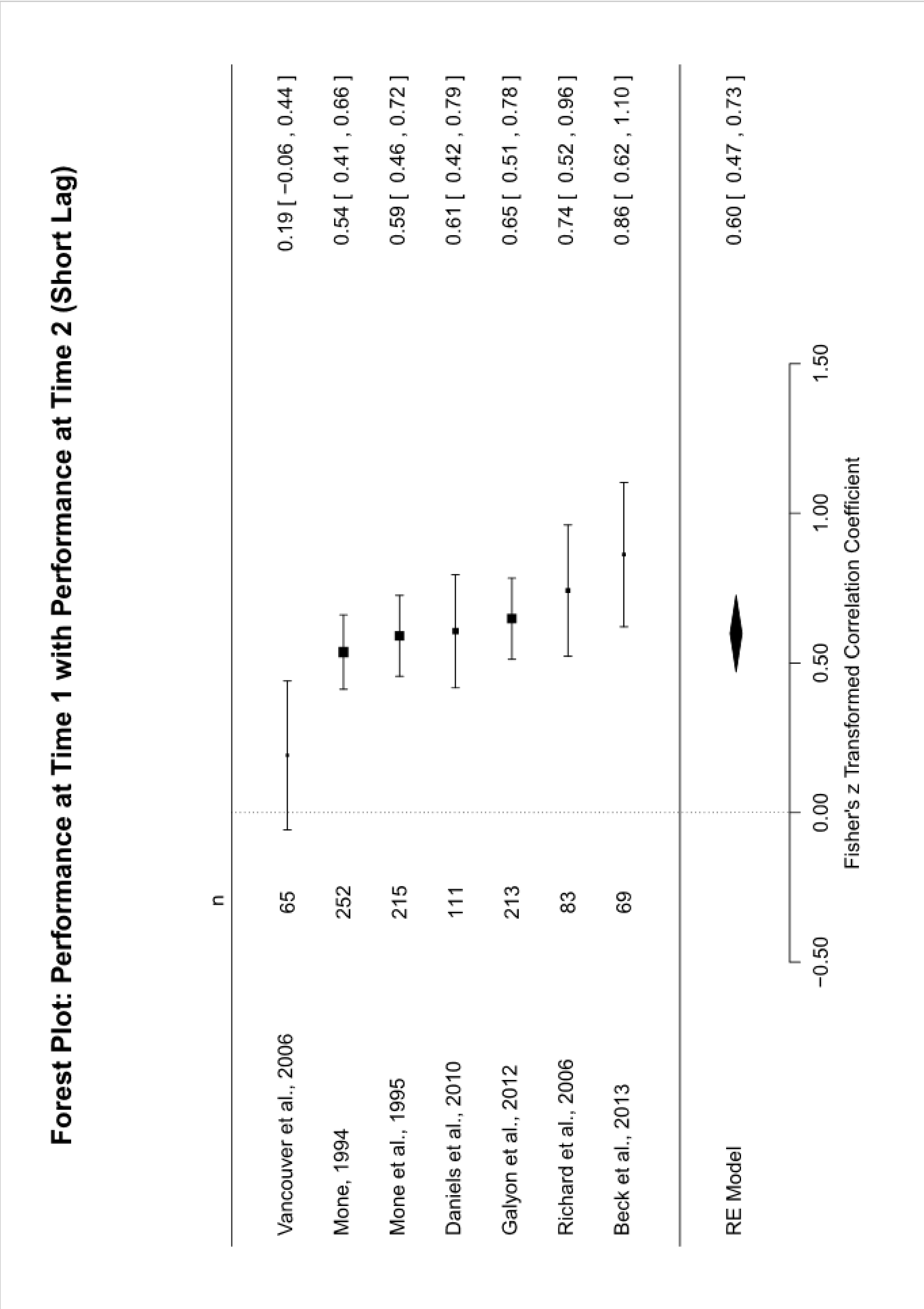


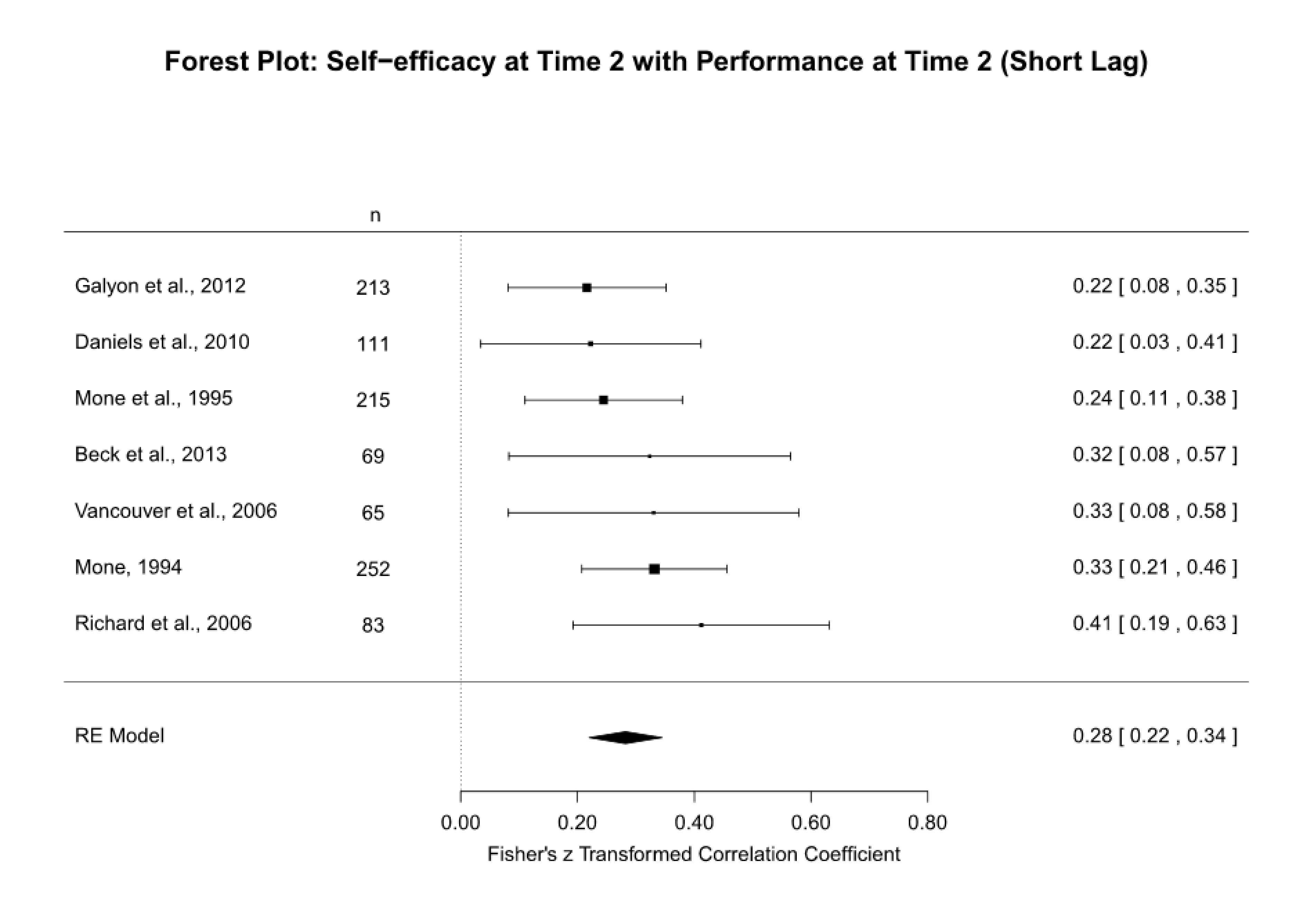
Forest Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2 (Short Lag)
n
Galyon et al., 2012
213
0.61 [0.48 , 0.75]
Vancouver et al., 2006
65
0.74 [0.49 , 0.99]
Mone et al., 1995
215
0.78 [0.64 , 0.91]
Mone, 1994
252
0.81 [0.69 , 0.93]
Daniels et al., 2010
111
0.82 [0.63 , 1.01]
Beck et al., 2013
69
0.94 [0.70 , 1.18]
Richard et al., 2006
83
1.00 [0.78 , 1.22]
RE Model
0.80 [0.70 , 0.89]
0.40
0.60
0.80
1.00
1.20
1.40
Fisher's z Transformed Correlation Coefficient

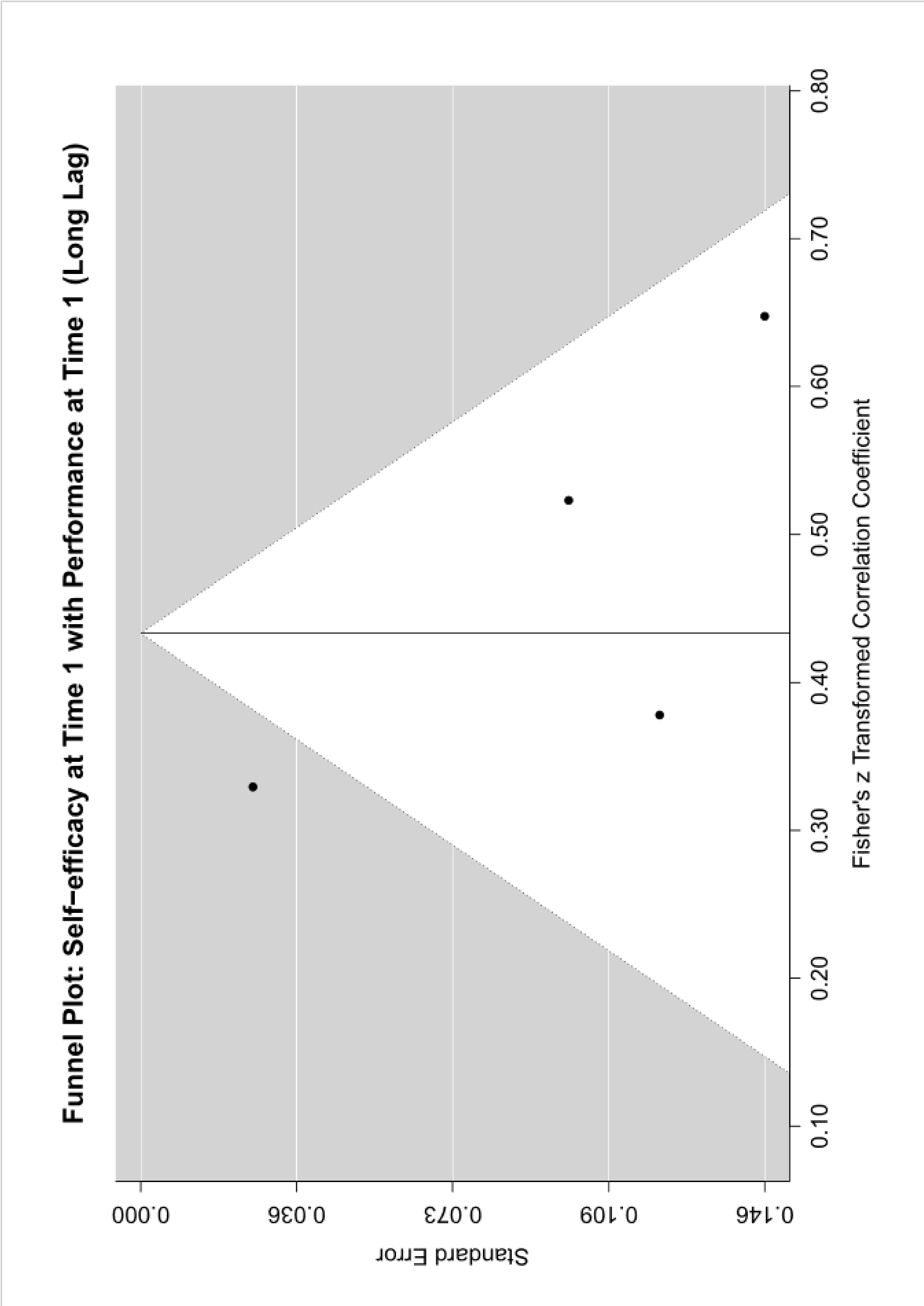

Forest Plot: Self−efficacy at Time 1 with Performance at Time 2 (Short Lag)
n
Galyon et al., 2012
213
0.03 [−0.11 , 0.16]
Mone et al., 1995
215
0.22 [0.09 , 0.36]
Daniels et al., 2010
111
0.23 [0.04 , 0.42]
Vancouver et al., 2006
65
0.26 [0.01 , 0.51]
Mone, 1994
252
0.30 [0.17 , 0.42]
Beck et al., 2013
69
0.30 [0.06 , 0.54]
Richard et al., 2006
83
0.35 [0.13 , 0.57]
RE Model
0.23 [0.14 , 0.32]
−0.20
0.00
0.20
0.40
0.60
Fisher's z Transformed Correlation Coefficient

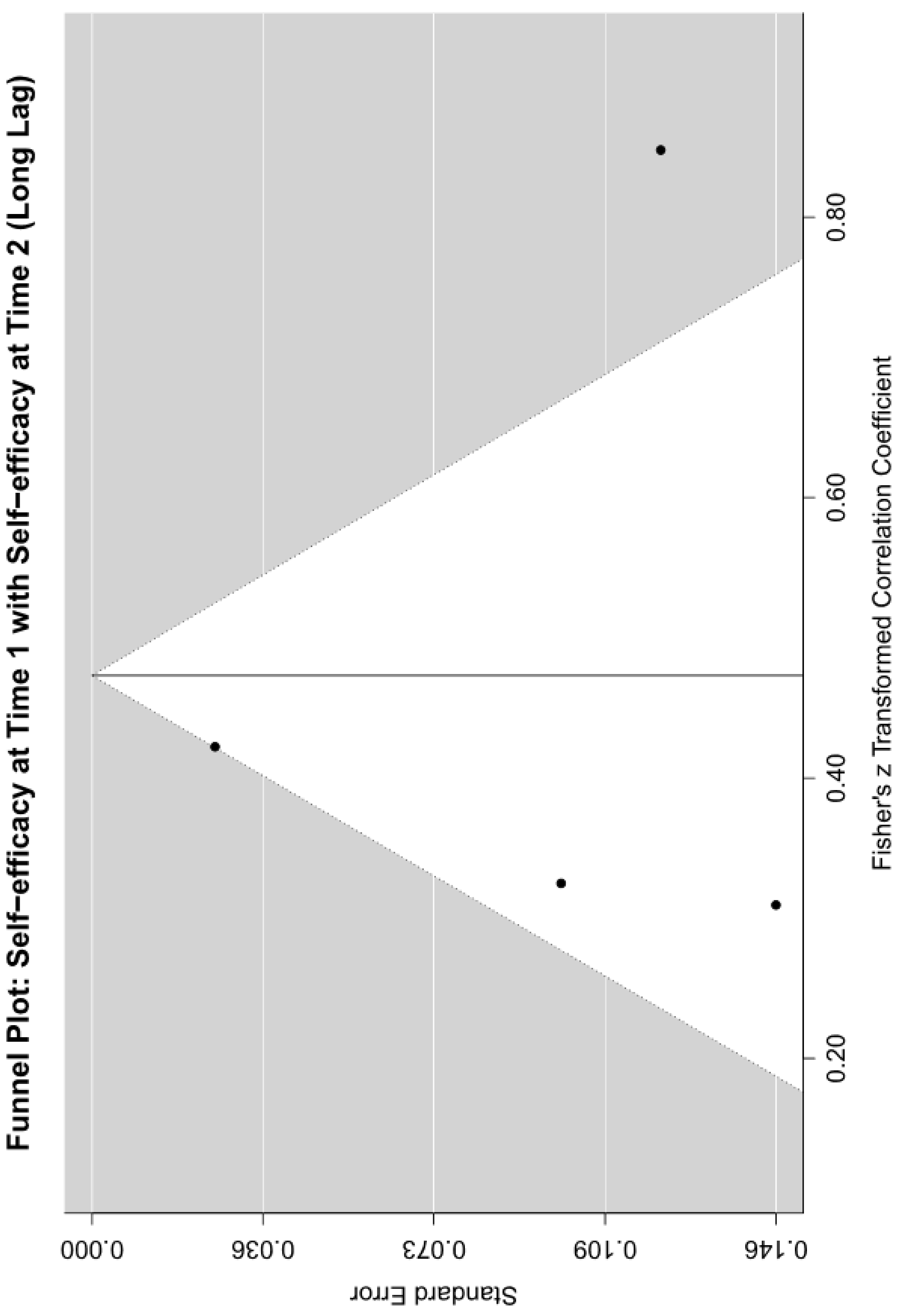


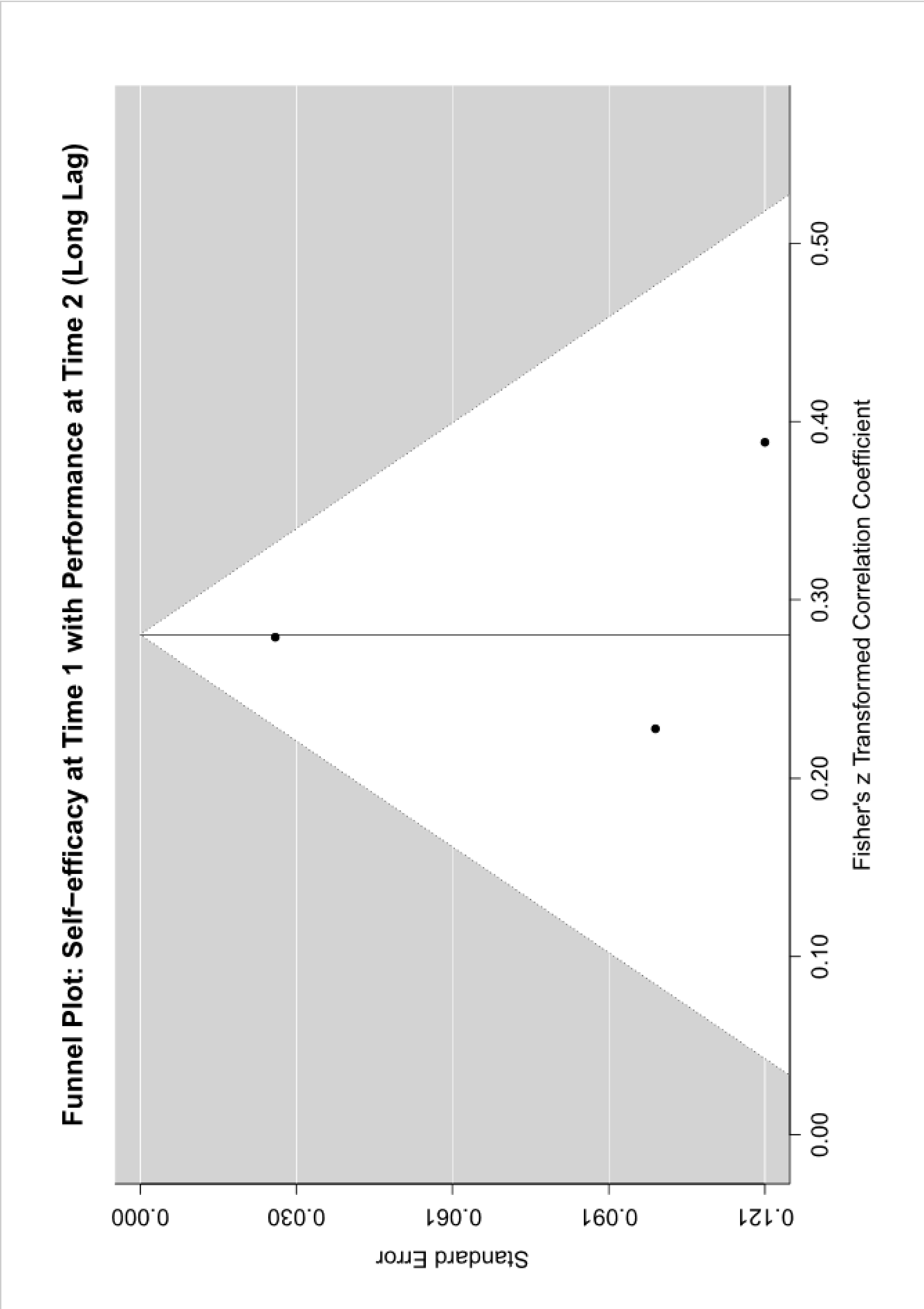
Forest Plot: Self–efficacy at Time 2 with Performance at Time 1 (Short Lag)
n
Galyon et al., 2012 213 0.24 [0.11 , 0.38]
Vancouver et al., 2006 65 0.43 [0.18 , 0.68]
Daniels et al., 2010 111 0.46 [0.27 , 0.65]
Mone et al., 1995 215 0.48 [0.35 , 0.62]
Mone, 1994 252 0.52 [0.40 , 0.65]
Richard et al., 2006 83 0.52 [0.30 , 0.74]
Beck et al., 2013 69 0.53 [0.29 , 0.77]
RE Model 0.45 [0.36 , 0.54]
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient


Forest Plot: Performance at Time 1 with Performance at Time 2 (Short Lag)
n
Vancouver et al., 2006
65
0.19 [−0.06 , 0.44]
Mone, 1994
252
0.54 [0.41 , 0.66]
Mone et al., 1995
215
0.59 [0.46 , 0.72]
Daniels et al., 2010
111
0.61 [0.42 , 0.79]
Galyon et al., 2012
213
0.65 [0.51 , 0.78]
Richard et al., 2006
83
0.74 [0.52 , 0.96]
Beck et al., 2013
69
0.86 [0.62 , 1.10]
RE Model
0.60 [0.47 , 0.73]
−0.50
0.00
0.50
1.00
1.50
Fisher's z Transformed Correlation Coefficient

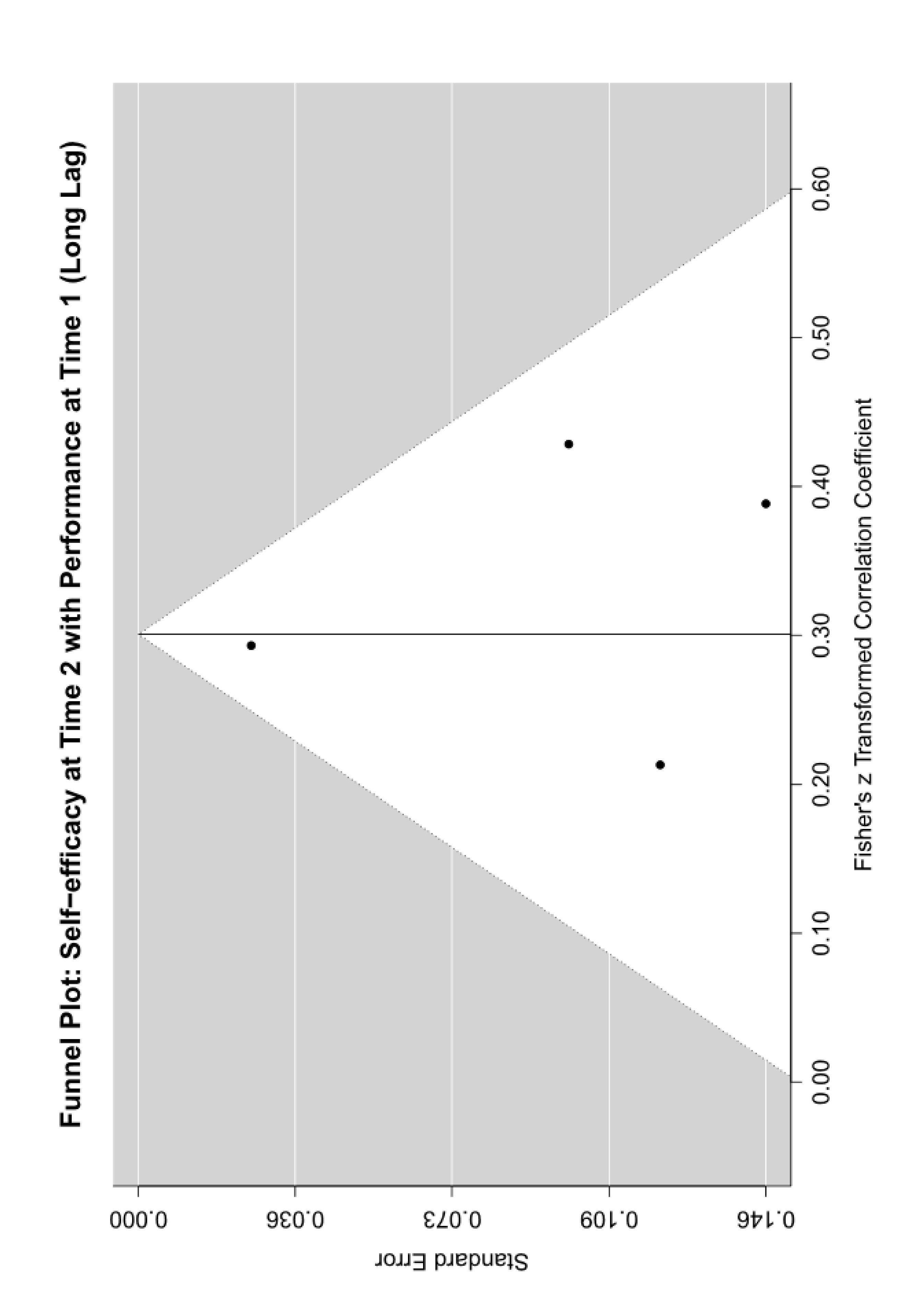

Forest Plot: Self−efficacy at Time 2 with Performance at Time 2 (Short Lag)
n
Galyon et al., 2012
213
0.22 [0.08 , 0.35]
Daniels et al., 2010
111
0.22 [0.03 , 0.41]
Mone et al., 1995
215
0.24 [0.11 , 0.38]
Beck et al., 2013
69
0.32 [0.08 , 0.57]
Vancouver et al., 2006
65
0.33 [0.08 , 0.58]
Mone, 1994
252
0.33 [0.21 , 0.46]
Richard et al., 2006
83
0.41 [0.19 , 0.63]
RE Model
0.28 [0.22 , 0.34]
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient

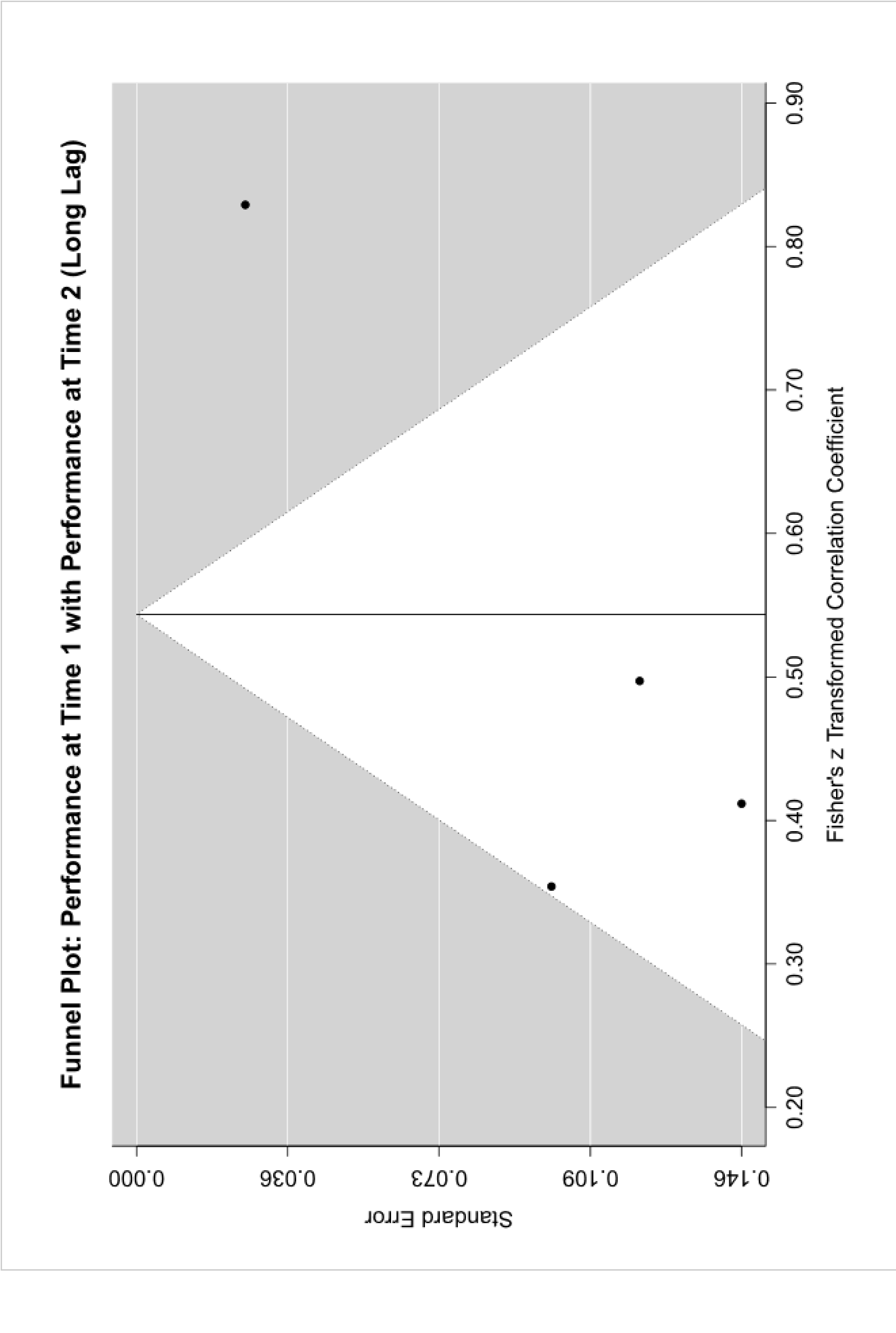

Funnel Plot: Self−efficacy at Time 1 with Performance at Time 1 (Long Lag)
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80

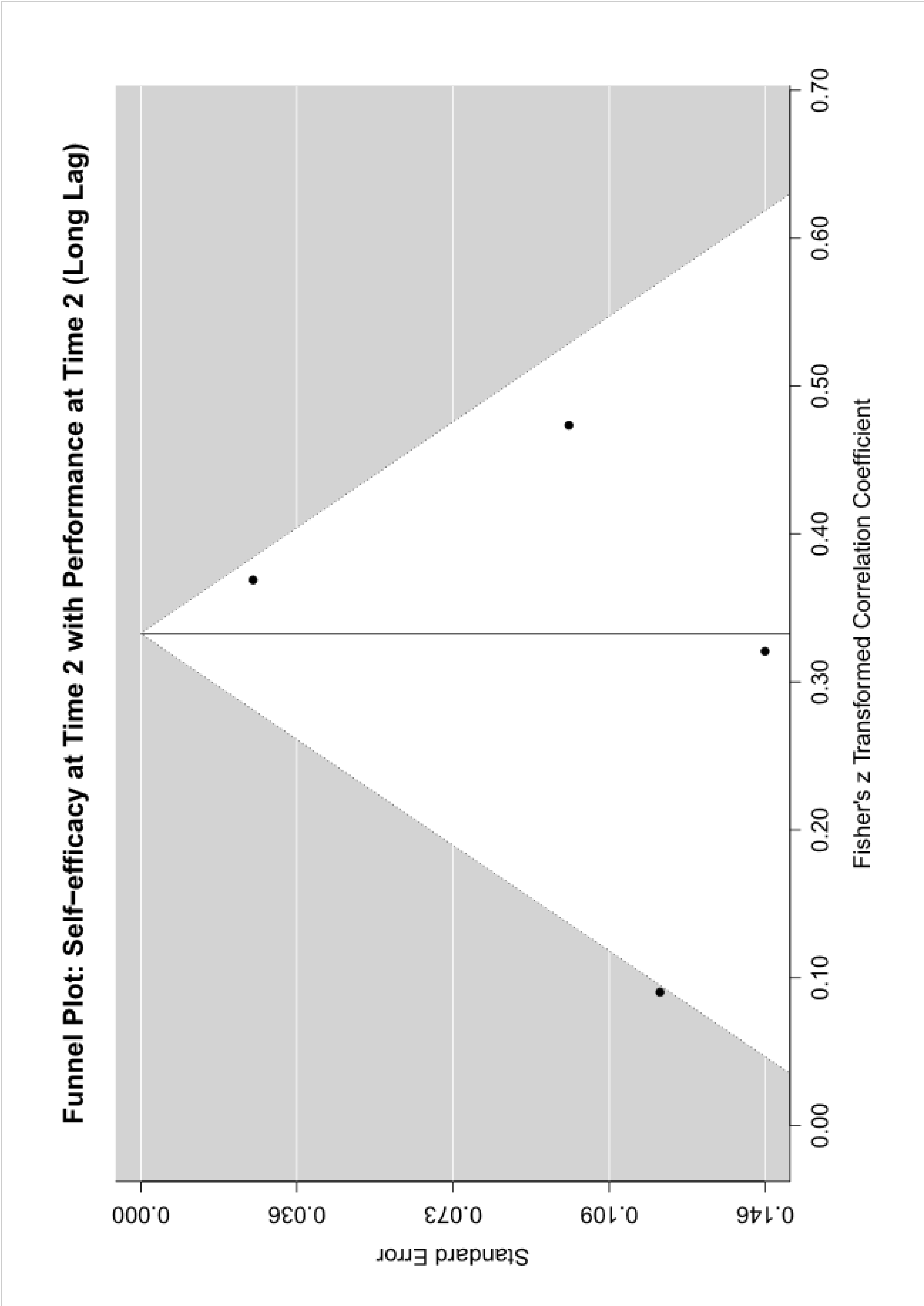

Funnel Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2 (Long Lag)
Standard Error
0.000
0.036
0.073
0.109
0.146
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient



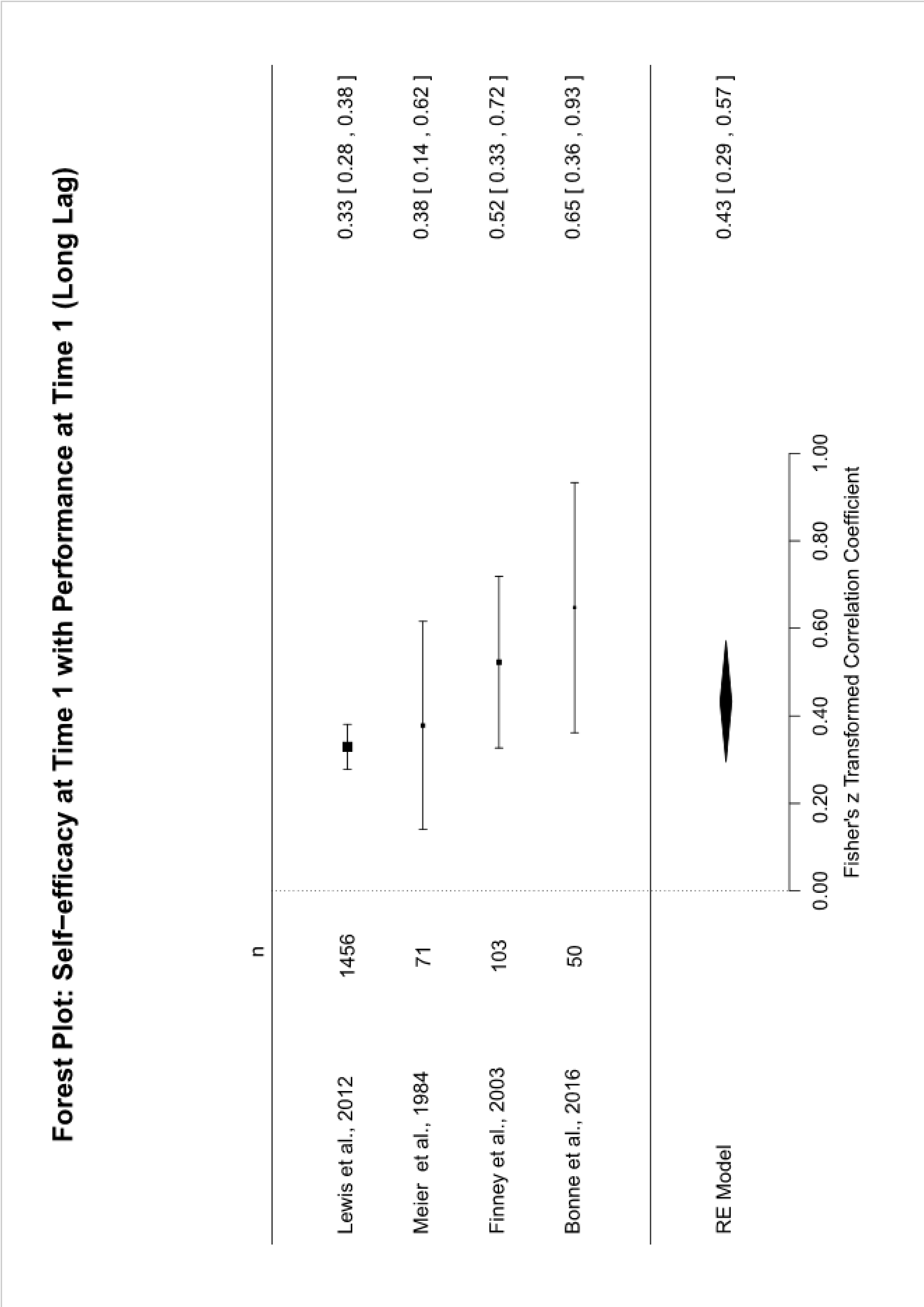
Funnel Plot: Self−efficacy at Time 1 with Performance at Time 2 (Long Lag)
Standard Error
0.000
0.030
0.061
0.091
0.121
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50

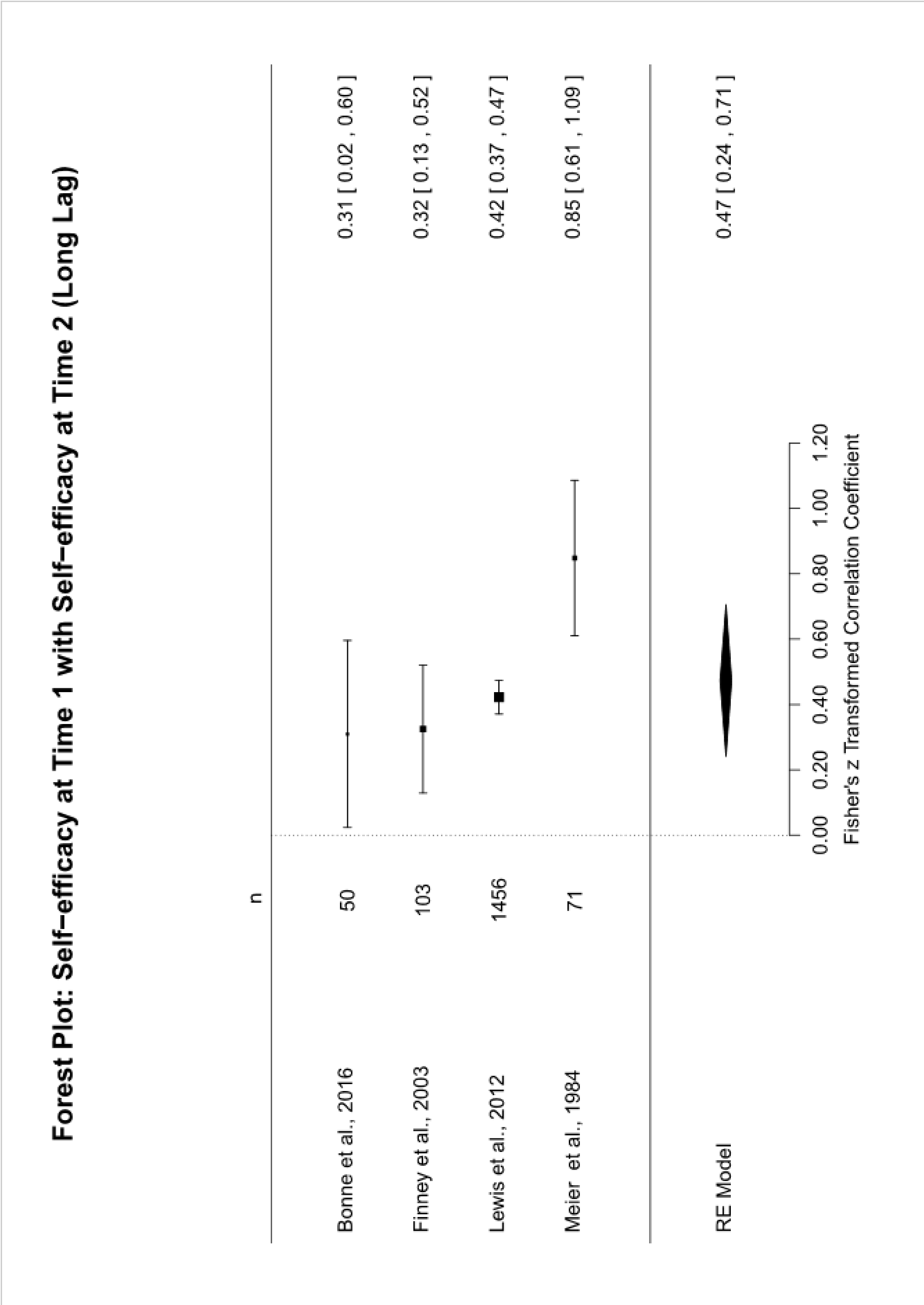

Funnel Plot: Self−efficacy at Time 2 with Performance at Time 1 (Long Lag)
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50
0.60

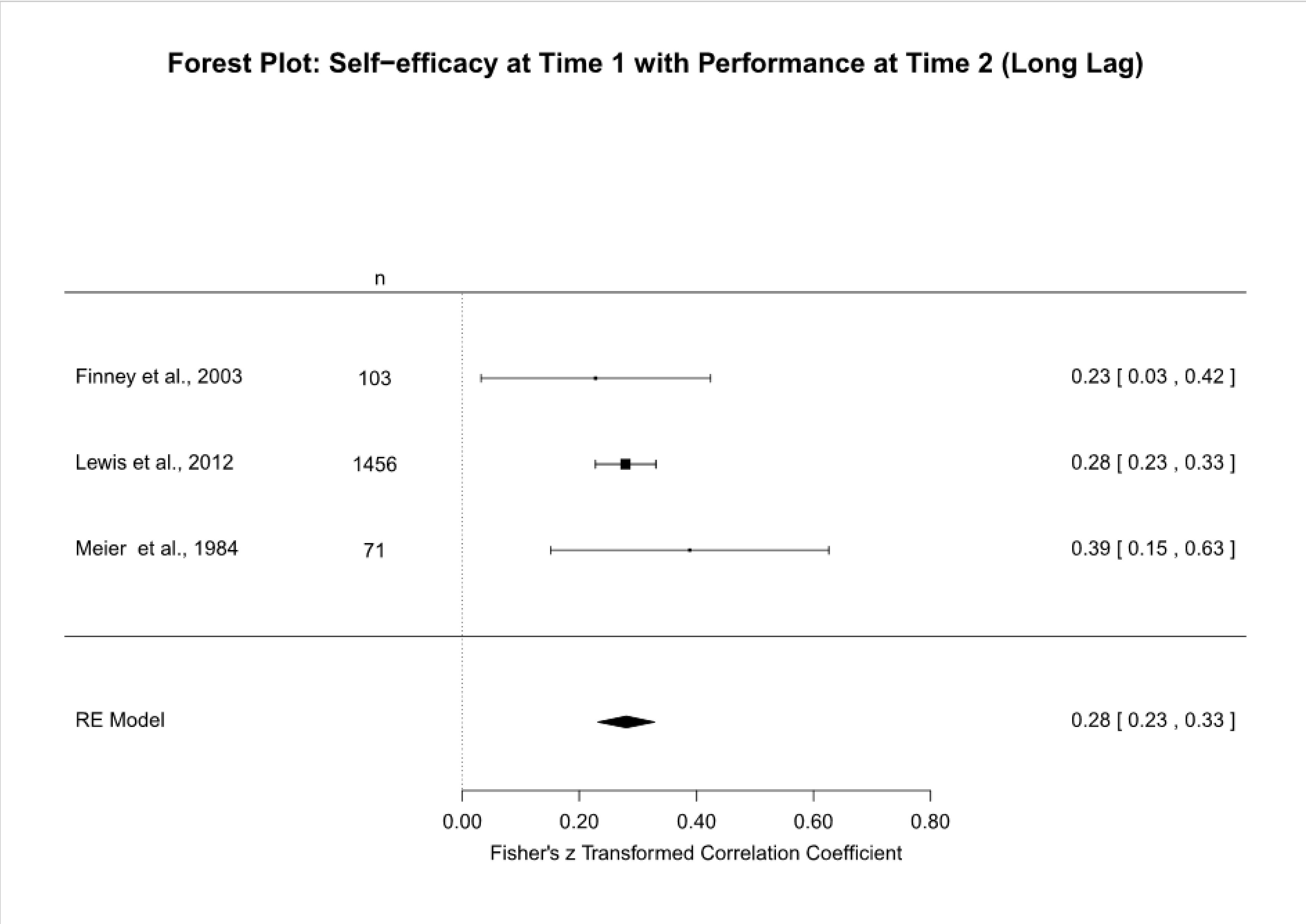

Funnel Plot: Performance at Time 1 with Performance at Time 2 (Long Lag)
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90

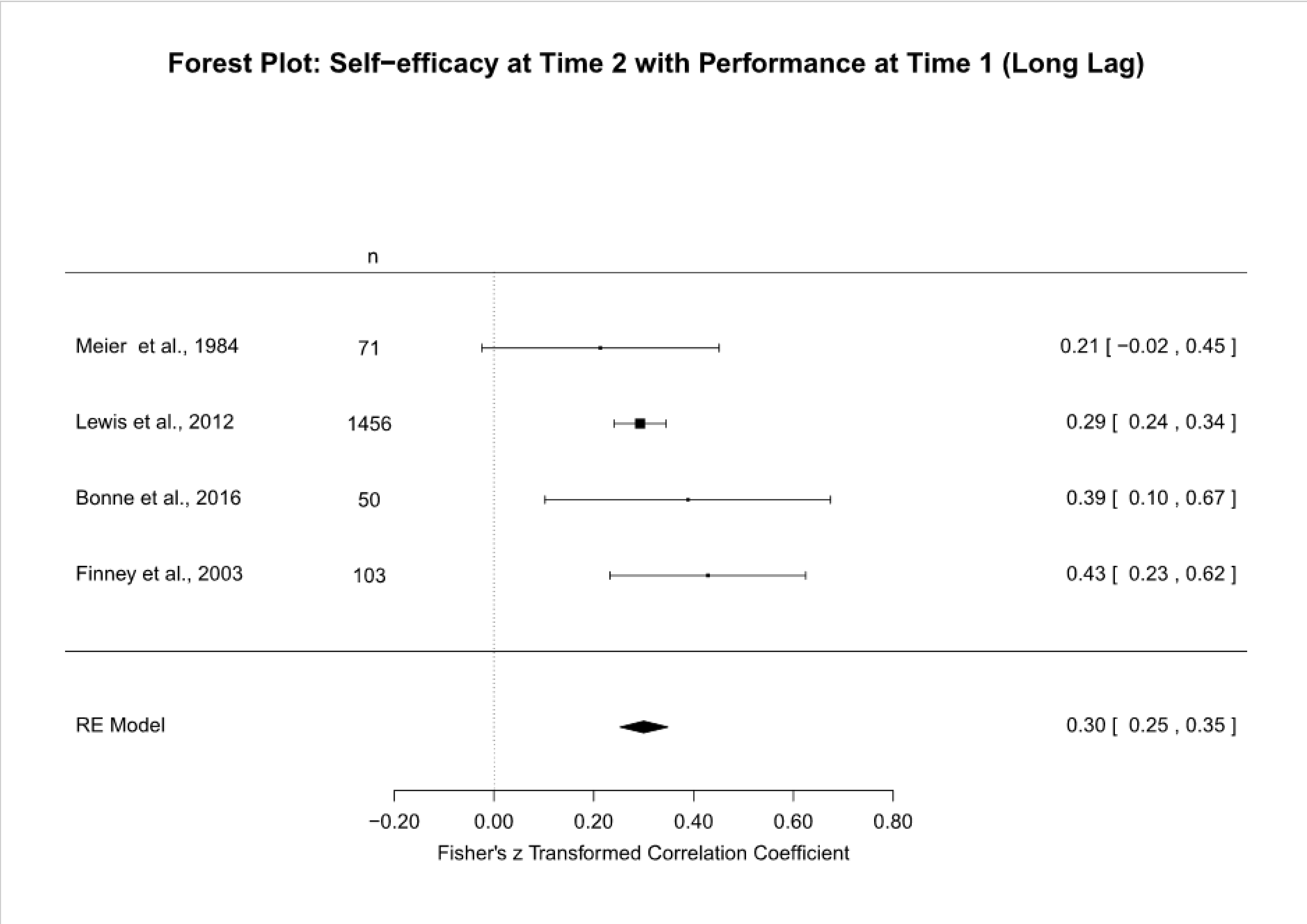


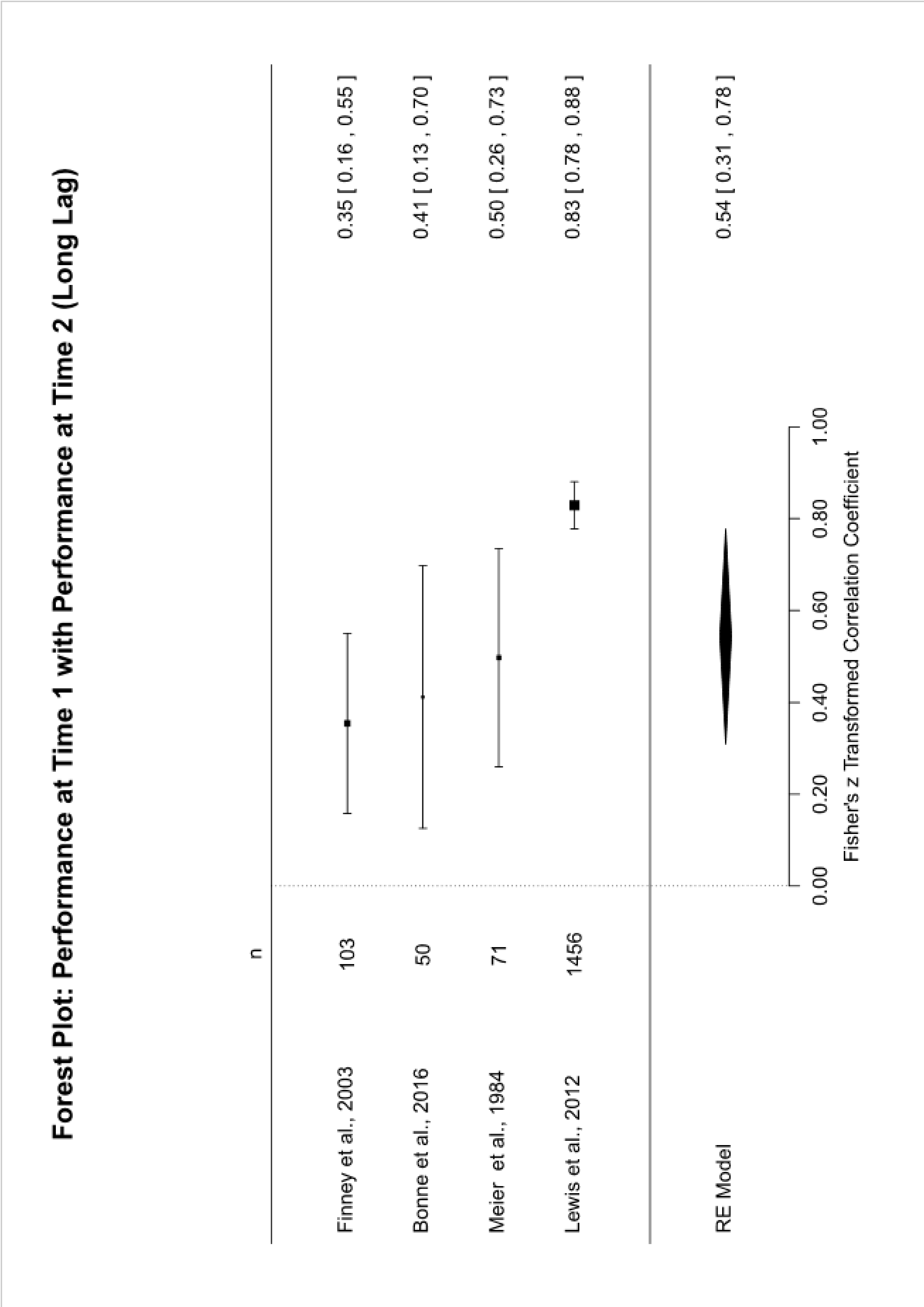
Funnel Plot: Self−efficacy at Time 2 with Performance at Time 2 (Long Lag)
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70

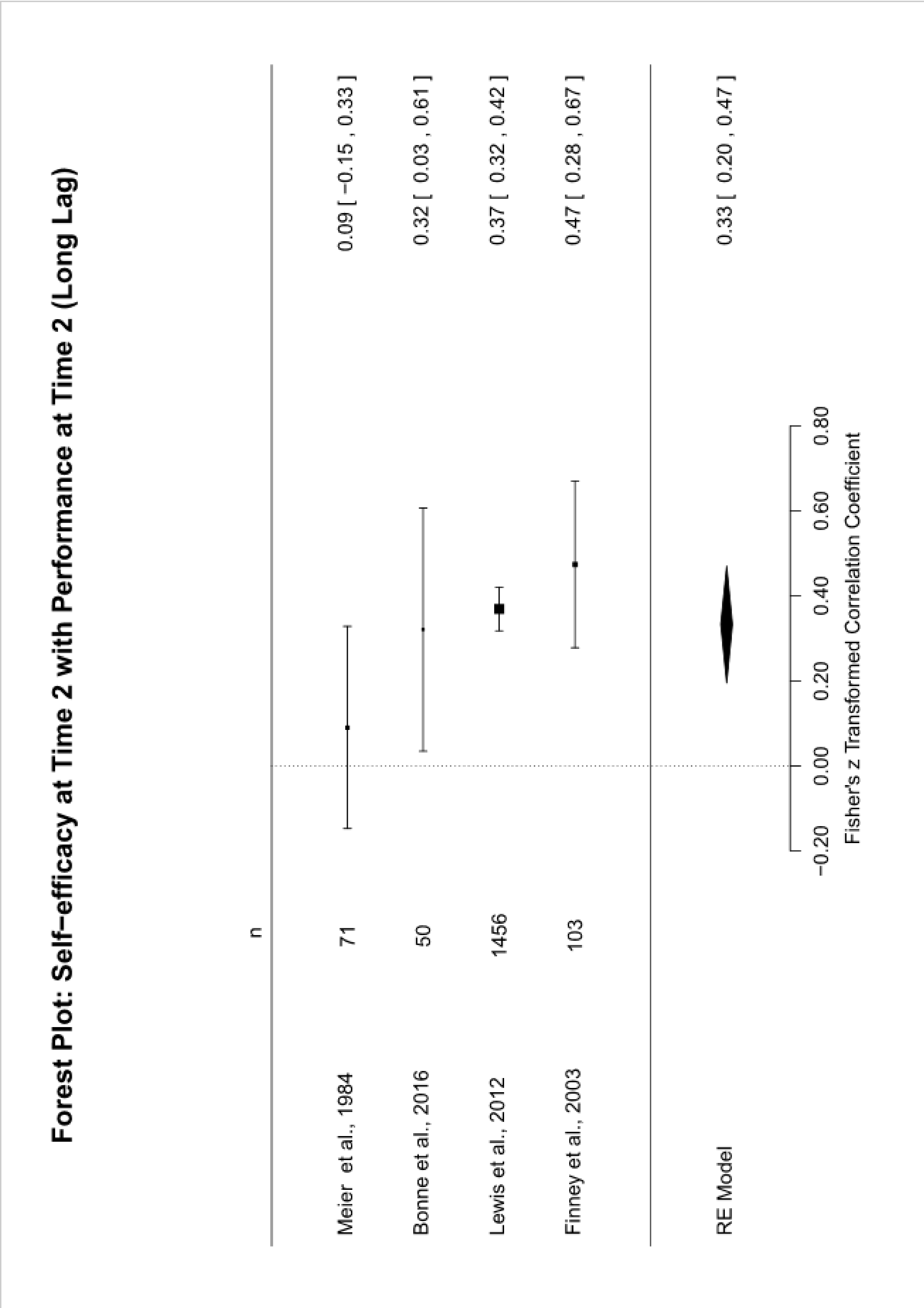

Forest Plot: Self−efficacy at Time 1 with Performance at Time 1 (Long Lag)
n
Lewis et al., 2012
1456
0.33 [0.28 , 0.38]
Meier et al., 1984
71
0.38 [0.14 , 0.62]
Finney et al., 2003
103
0.52 [0.33 , 0.72]
Bonne et al., 2016
50
0.65 [0.36 , 0.93]
RE Model
0.43 [0.29 , 0.57]
0.00
0.20
0.40
0.60
0.80
1.00
Fisher's z Transformed Correlation Coefficient

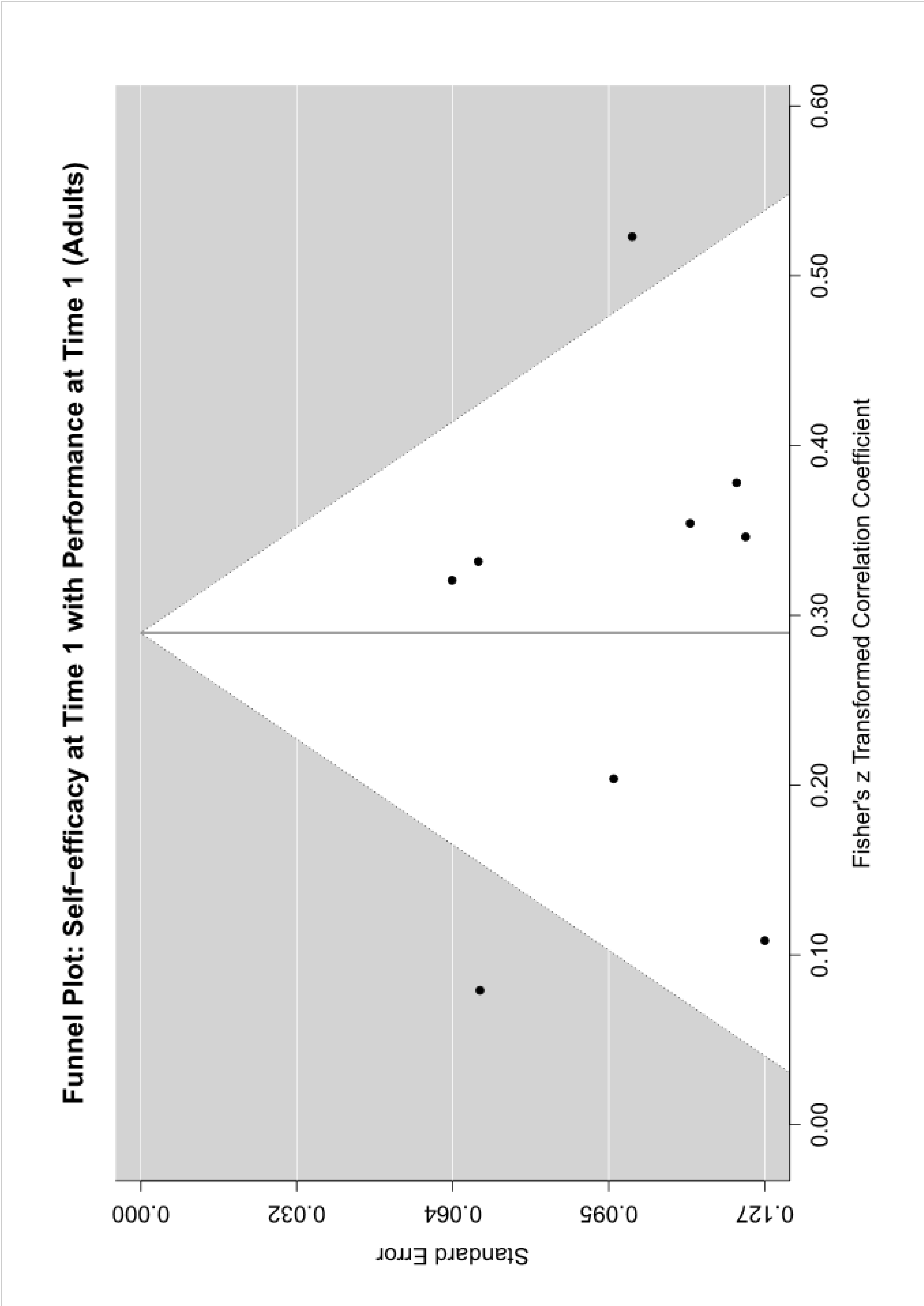

Forest Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2 (Long Lag)
n
Bonne et al., 2016
50
0.31 [0.02 , 0.60]
Finney et al., 2003
103
0.32 [0.13 , 0.52]
Lewis et al., 2012
1456
0.42 [0.37 , 0.47]
Meier et al., 1984
71
0.85 [0.61 , 1.09]
RE Model
0.47 [0.24 , 0.71]
0.00
0.20
0.40
0.60
0.80
1.00
1.20
Fisher's z Transformed Correlation Coefficient


Forest Plot: Self−efficacy at Time 1 with Performance at Time 2 (Long Lag)
n
Finney et al., 2003
103
0.23 [0.03 , 0.42]
Lewis et al., 2012
1456
0.28 [0.23 , 0.33]
Meier et al., 1984
71
0.39 [0.15 , 0.63]
RE Model
0.28 [0.23 , 0.33]
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient

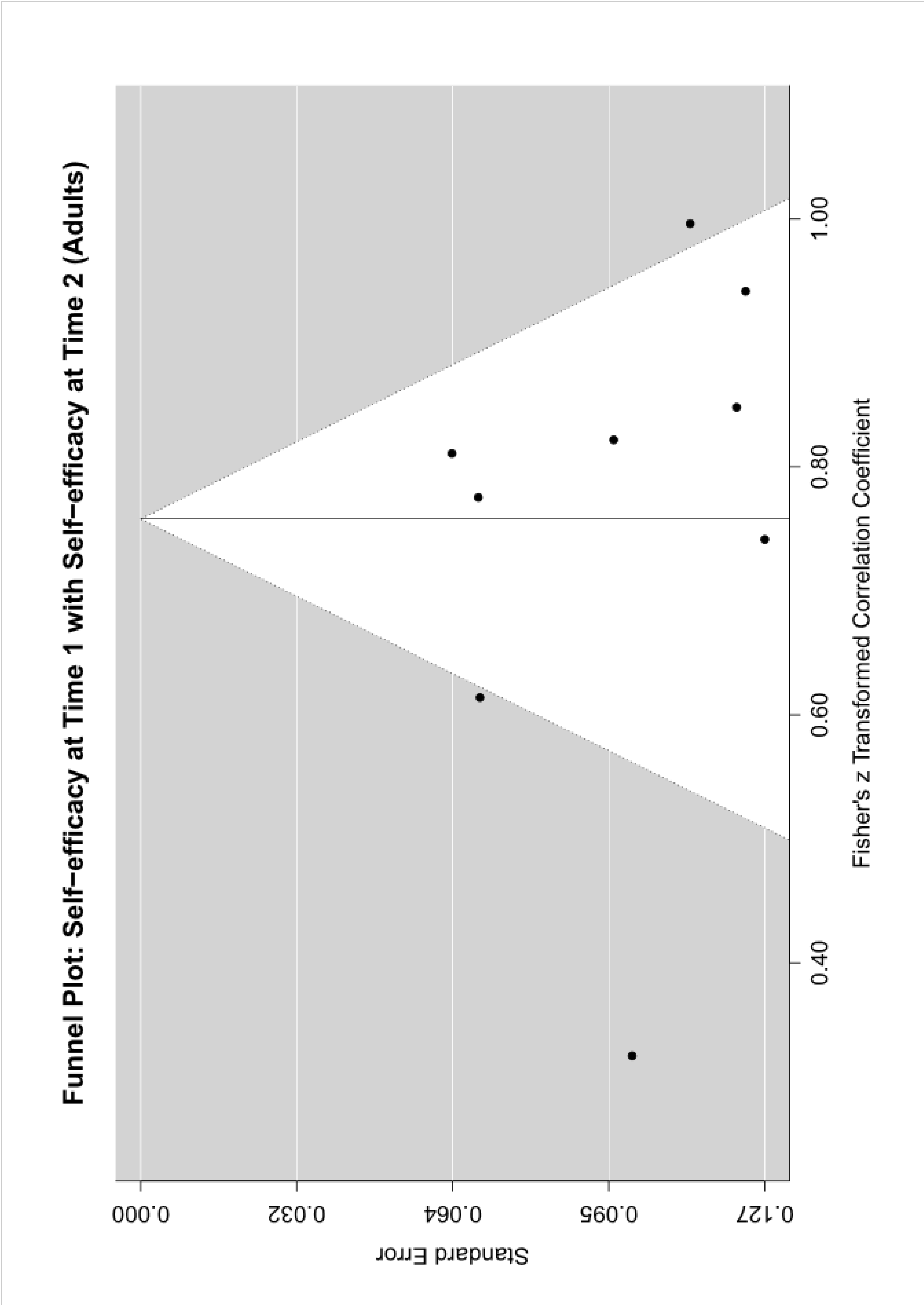

Forest Plot: Self−efficacy at Time 2 with Performance at Time 1 (Long Lag)
n
Meier et al., 1984
71
0.21 [−0.02 , 0.45]
Lewis et al., 2012
1456
0.29 [0.24 , 0.34]
Bonne et al., 2016
50
0.39 [0.10 , 0.67]
Finney et al., 2003
103
0.43 [0.23 , 0.62]
RE Model
0.30 [0.25 , 0.35]
−0.20
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient


Forest Plot: Performance at Time 1 with Performance at Time 2 (Long Lag)
n
Finney et al., 2003
103
0.35 [0.16 , 0.55]
Bonne et al., 2016
50
0.41 [0.13 , 0.70]
Meier et al., 1984
71
0.50 [0.26 , 0.73]
Lewis et al., 2012
1456
0.83 [0.78 , 0.88]
RE Model
0.54 [0.31 , 0.78]
0.00
0.20
0.40
0.60
0.80
1.00
Fisher's z Transformed Correlation Coefficient

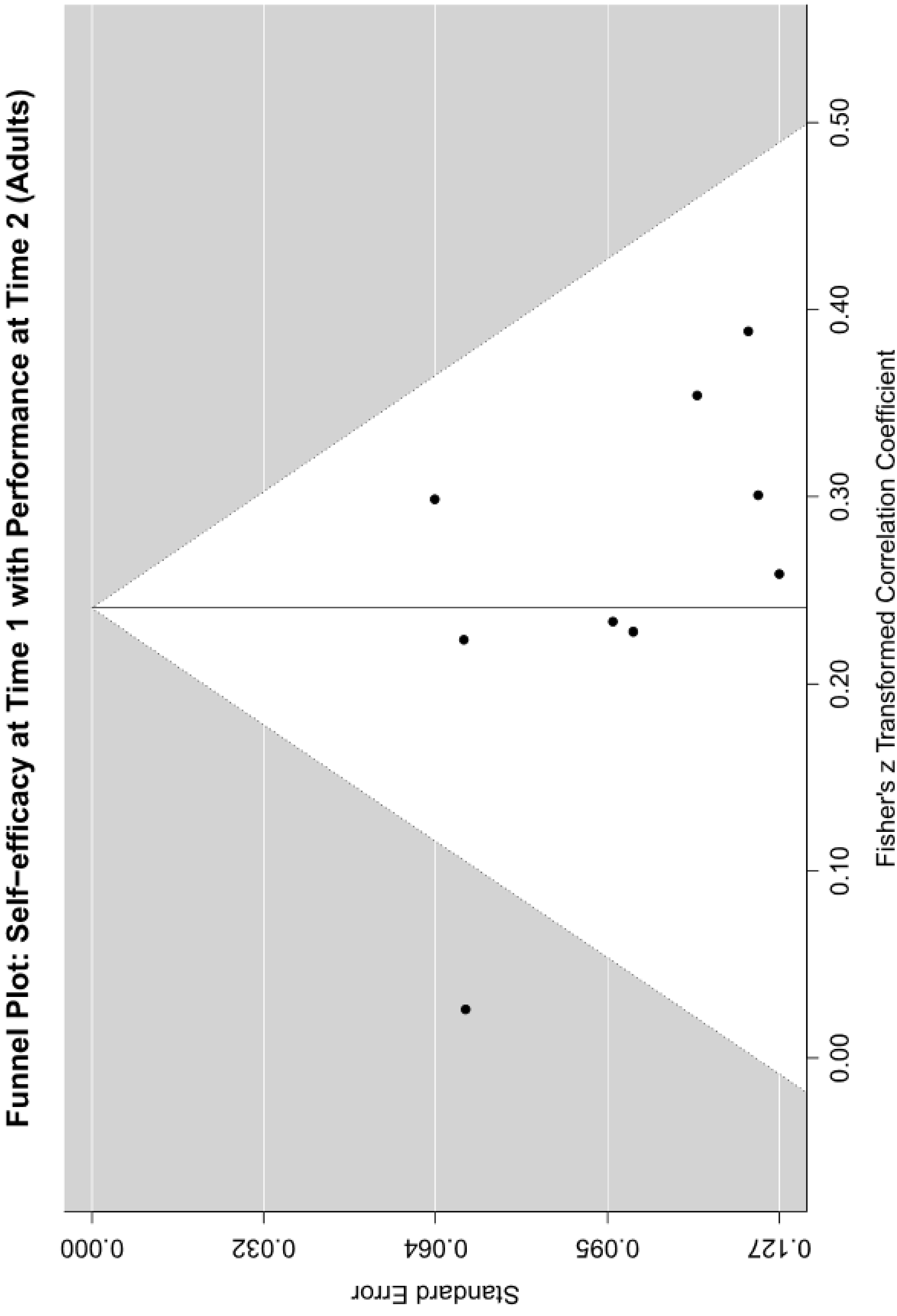

Forest Plot: Self−efficacy at Time 2 with Performance at Time 2 (Long Lag)
n
Meier et al., 1984
71
0.09 [−0.15 , 0.33]
Bonne et al., 2016
50
0.32 [0.03 , 0.61]
Lewis et al., 2012
1456
0.37 [0.32 , 0.42]
Finney et al., 2003
103
0.47 [0.28 , 0.67]
RE Model
0.33 [0.20 , 0.47]
−0.20
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient

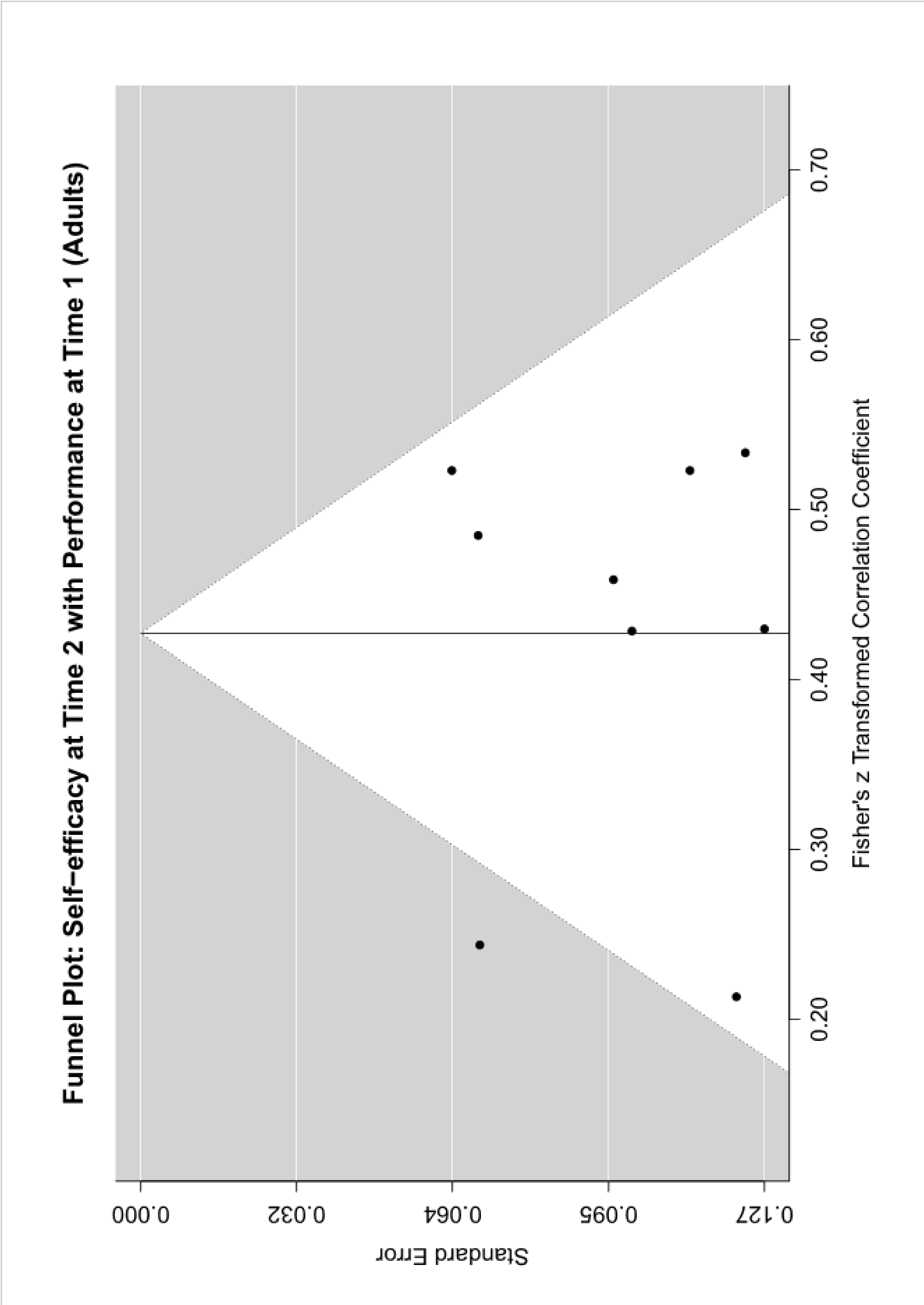


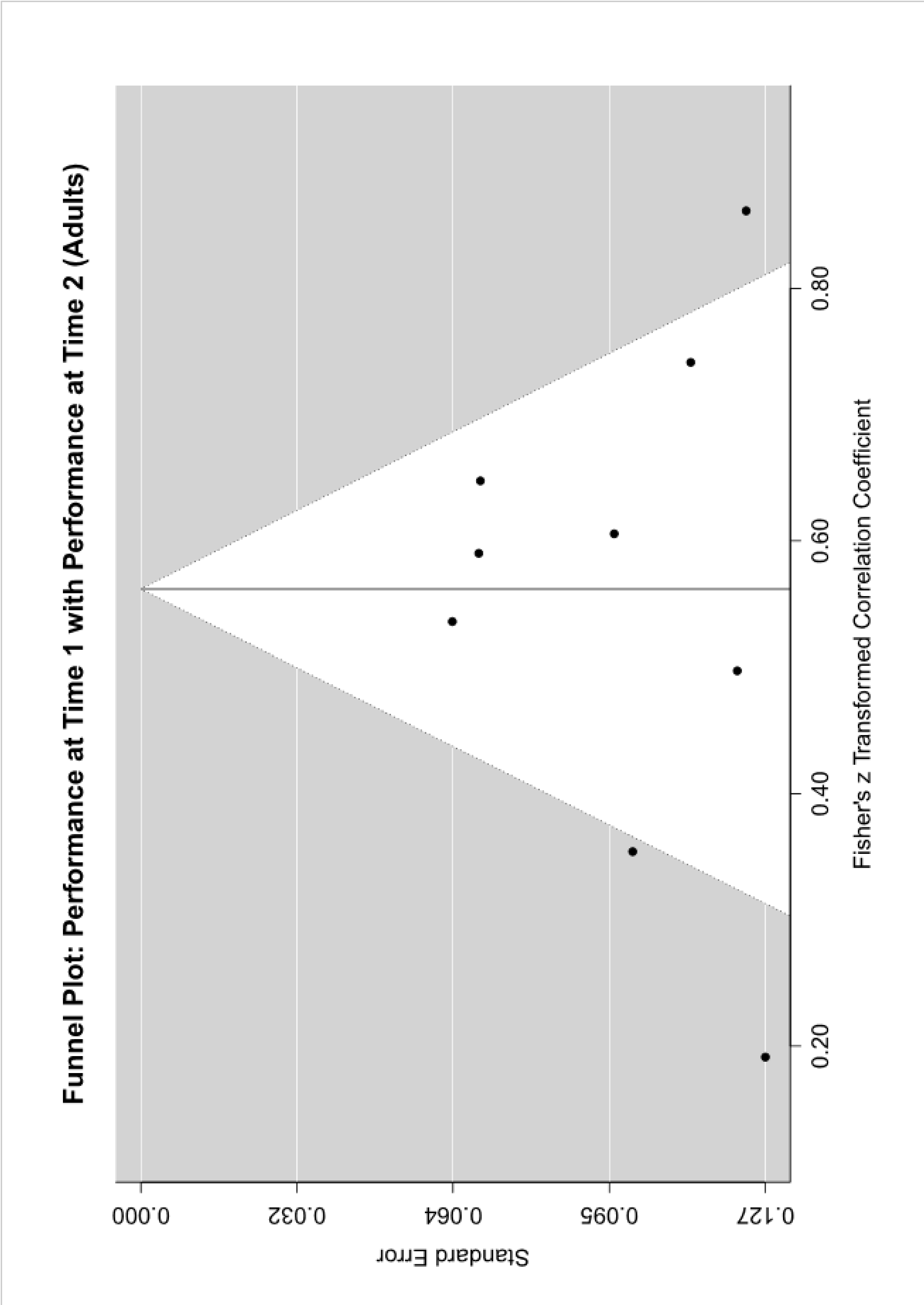
Funnel Plot: Self–efficacy at Time 1 with Performance at Time 1 (Adults)
Standard Error
0.000
0.032
0.064
0.095
0.127
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50
0.60

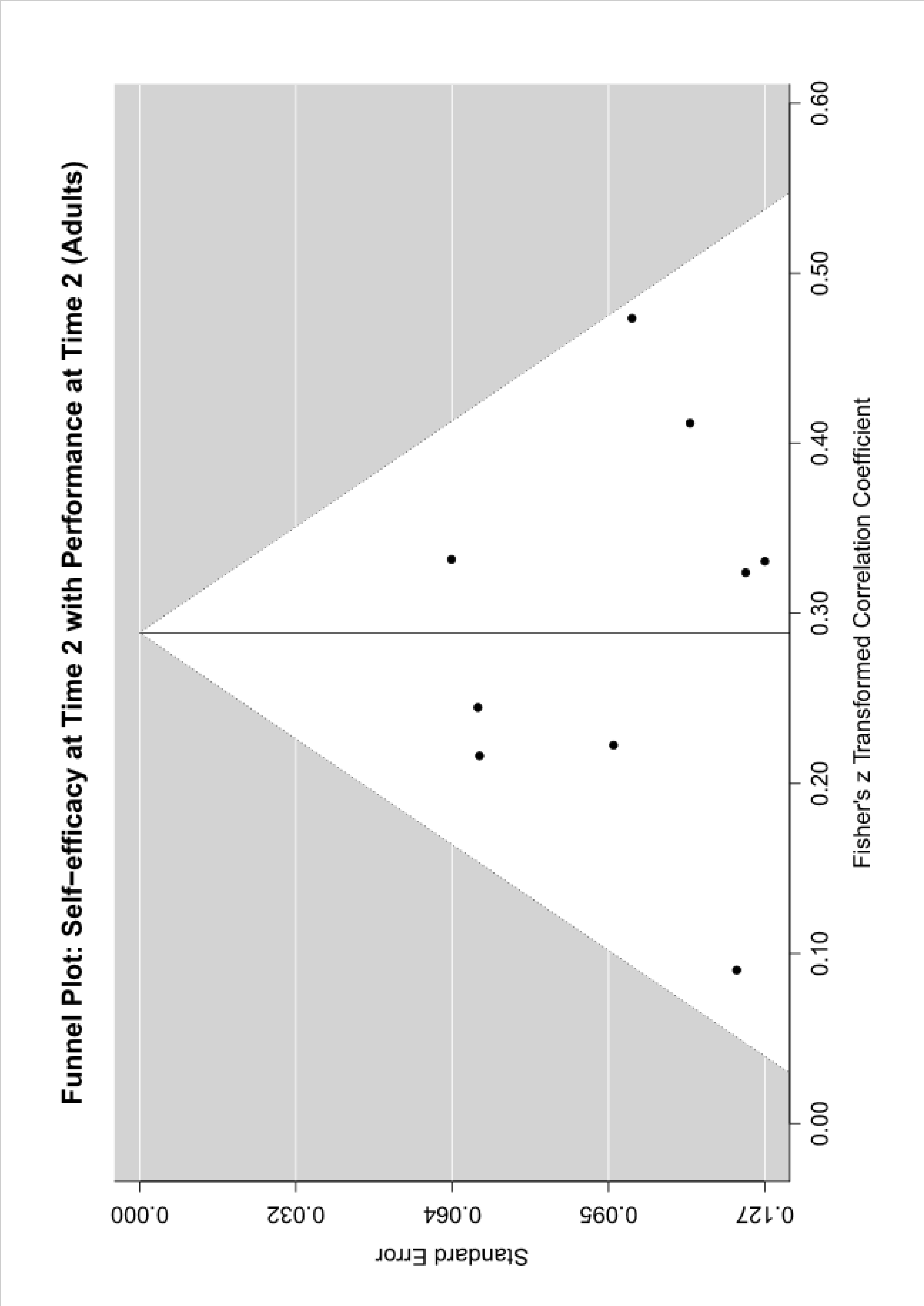


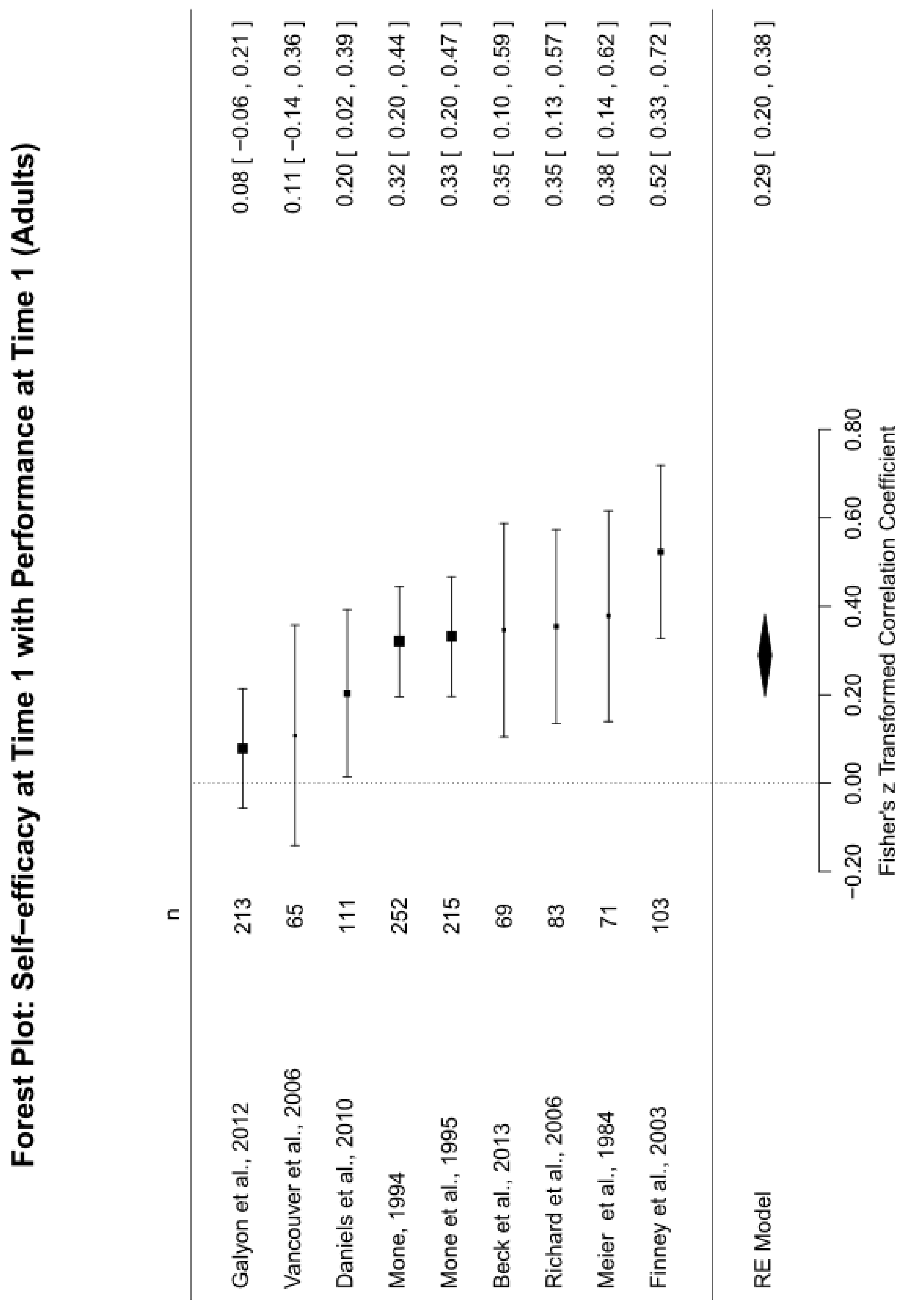
Funnel Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2 (Adults)
Standard Error
0.000
0.032
0.064
0.095
0.127
Fisher's z Transformed Correlation Coefficient
0.40
0.60
0.80
1.00

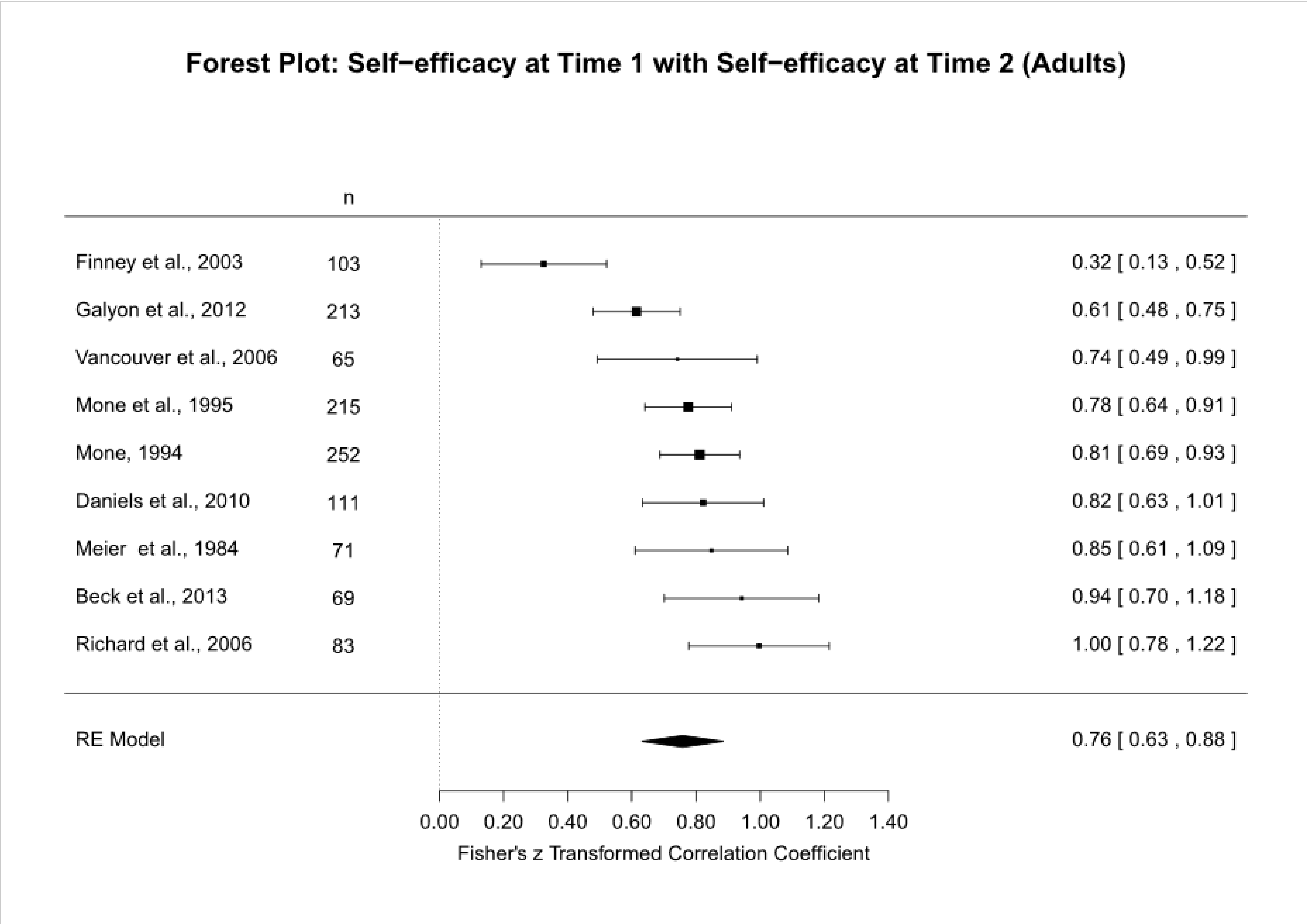

Funnel Plot: Self−efficacy at Time 1 with Performance at Time 2 (Adults)
Standard Error
0.000
0.032
0.064
0.095
0.127
0.00
0.10
0.20
0.30
0.40
0.50
Fisher's z Transformed Correlation Coefficient

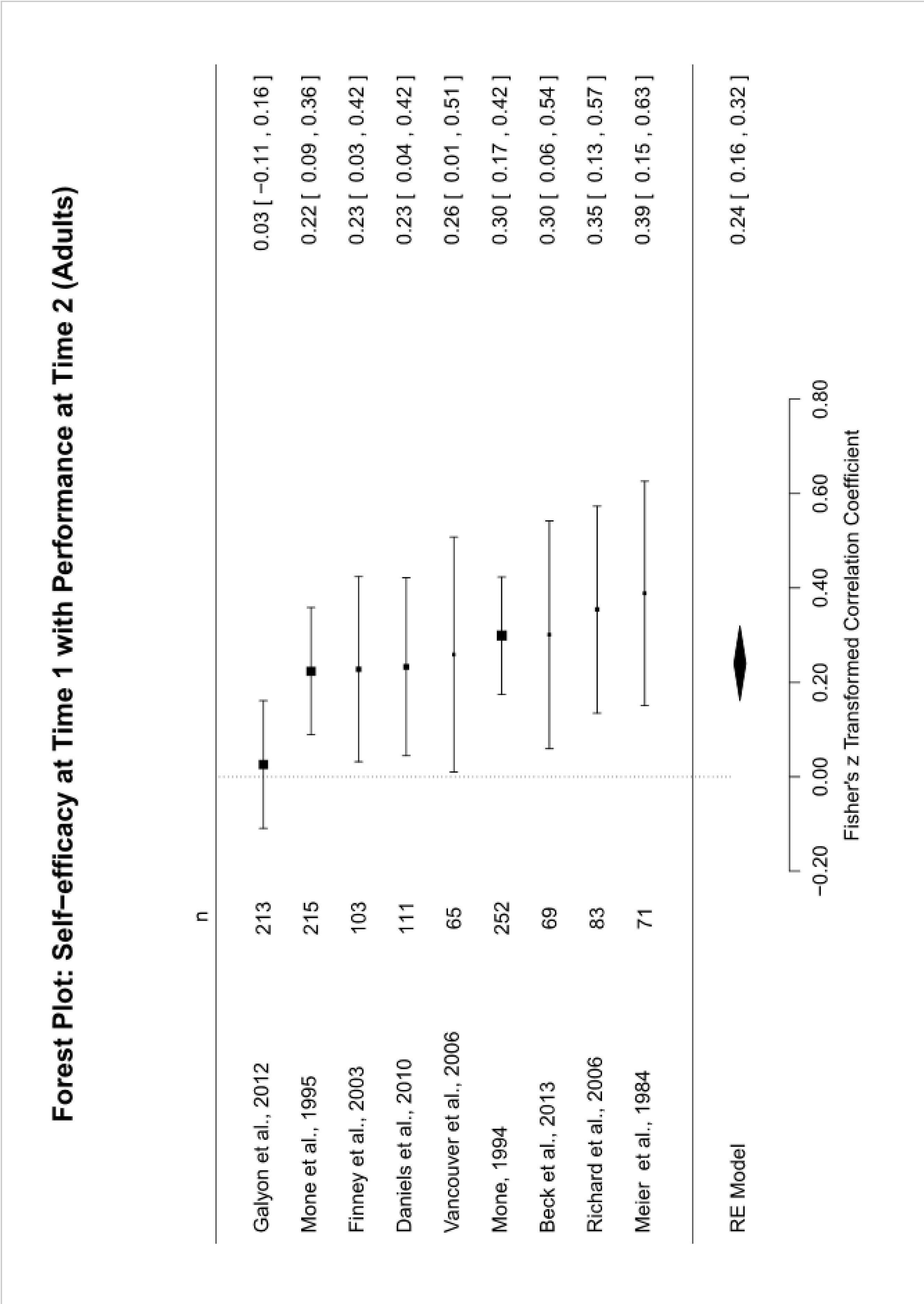

Funnel Plot: Self−efficacy at Time 2 with Performance at Time 1 (Adults)
Standard Error
0.000
0.032
0.064
0.095
0.127
Fisher's z Transformed Correlation Coefficient
0.20
0.30
0.40
0.50
0.60
0.70

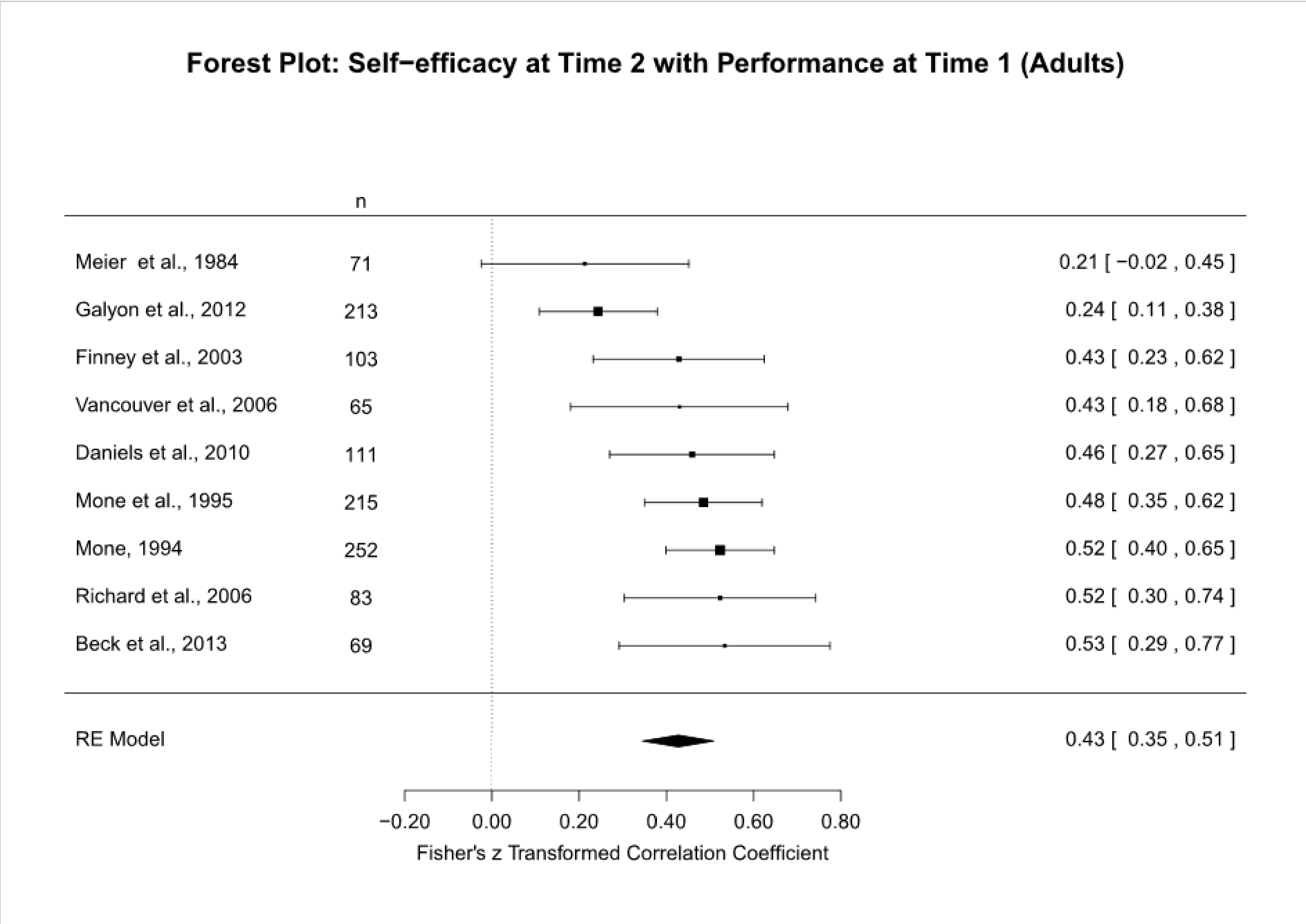

Funnel Plot: Performance at Time 1 with Performance at Time 2 (Adults)
Standard Error
0.000
0.032
0.064
0.095
0.127
Fisher's z Transformed Correlation Coefficient
0.20
0.40
0.60
0.80

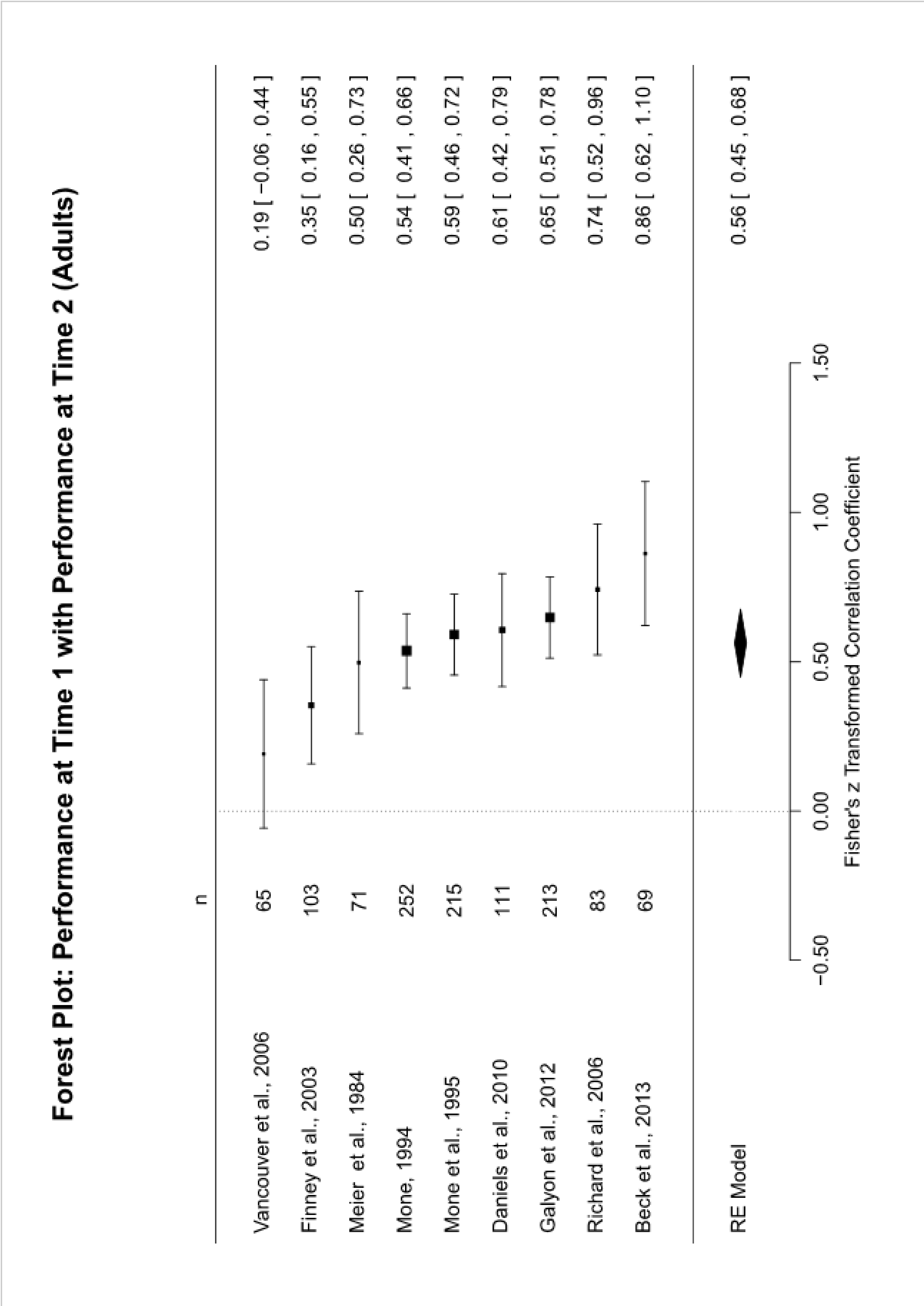

Funnel Plot: Self−efficacy at Time 2 with Performance at Time 2 (Adults)
Standard Error
0.000
0.032
0.064
0.095
0.127
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50
0.60


Forest Plot: Self–efficacy at Time 1 with Performance at Time 1 (Adults)
n
Galyon et al., 2012 213 0.08 [−0.06 , 0.21]
Vancouver et al., 2006 65 0.11 [−0.14 , 0.36]
Daniels et al., 2010 111 0.20 [0.02 , 0.39]
Mone, 1994 252 0.32 [0.20 , 0.44]
Mone et al., 1995 215 0.33 [0.20 , 0.47]
Beck et al., 2013 69 0.35 [0.10 , 0.59]
Richard et al., 2006 83 0.35 [0.13 , 0.57]
Meier et al., 1984 71 0.38 [0.14 , 0.62]
Finney et al., 2003 103 0.52 [0.33 , 0.72]
RE Model 0.29 [0.20 , 0.38]
−0.20 0.00 0.20 0.40 0.60 0.80
Fisher's z Transformed Correlation Coefficient

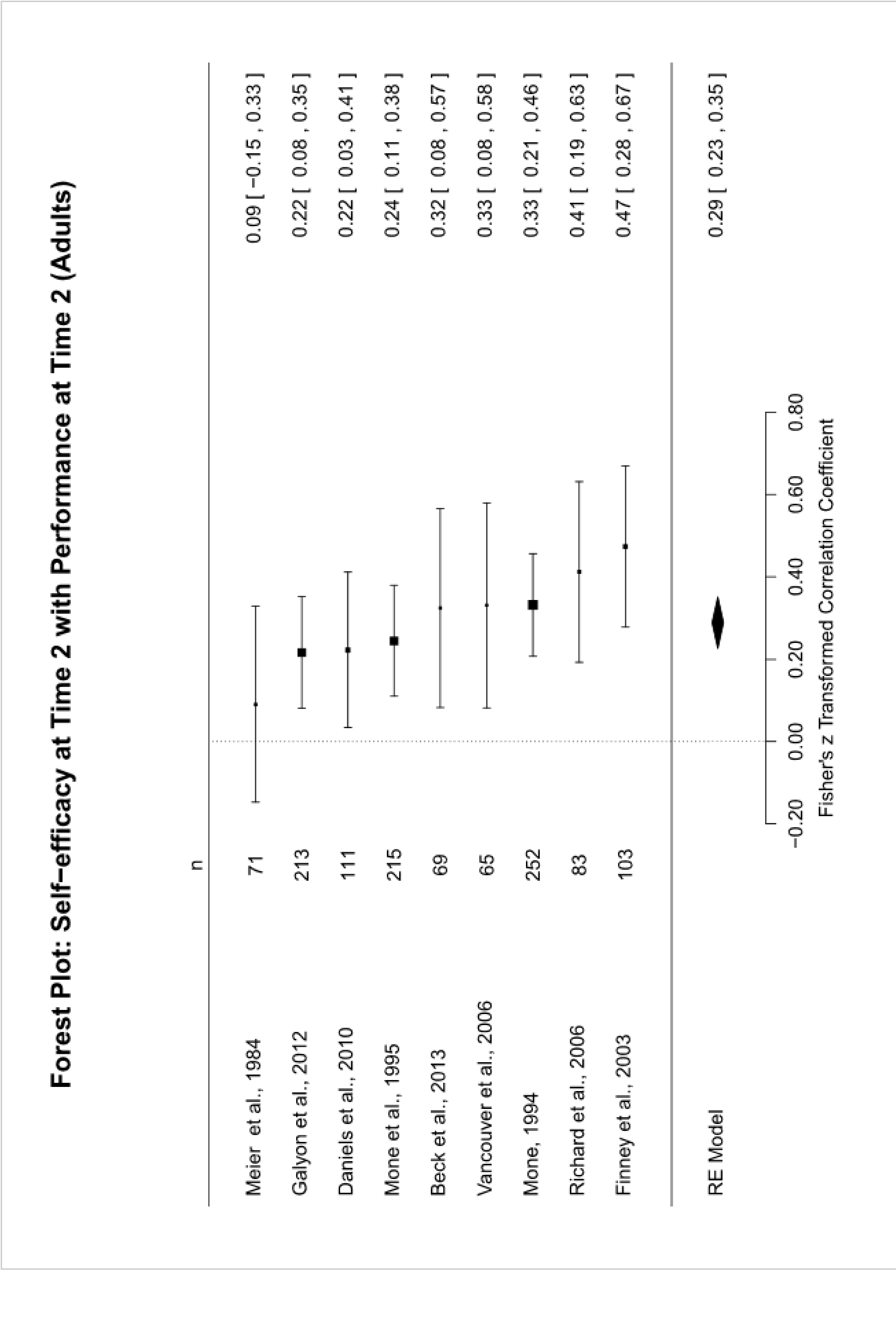

Forest Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2 (Adults)
n
Finney et al., 2003 103 0.32 [0.13 , 0.52]
Galyon et al., 2012 213 0.61 [0.48 , 0.75]
Vancouver et al., 2006 65 0.74 [0.49 , 0.99]
Mone et al., 1995 215 0.78 [0.64 , 0.91]
Mone, 1994 252 0.81 [0.69 , 0.93]
Daniels et al., 2010 111 0.82 [0.63 , 1.01]
Meier et al., 1984 71 0.85 [0.61 , 1.09]
Beck et al., 2013 69 0.94 [0.70 , 1.18]
Richard et al., 2006 83 1.00 [0.78 , 1.22]
RE Model 0.76 [0.63 , 0.88]
0.00 0.20 0.40 0.60 0.80 1.00 1.20 1.40
Fisher's z Transformed Correlation Coefficient

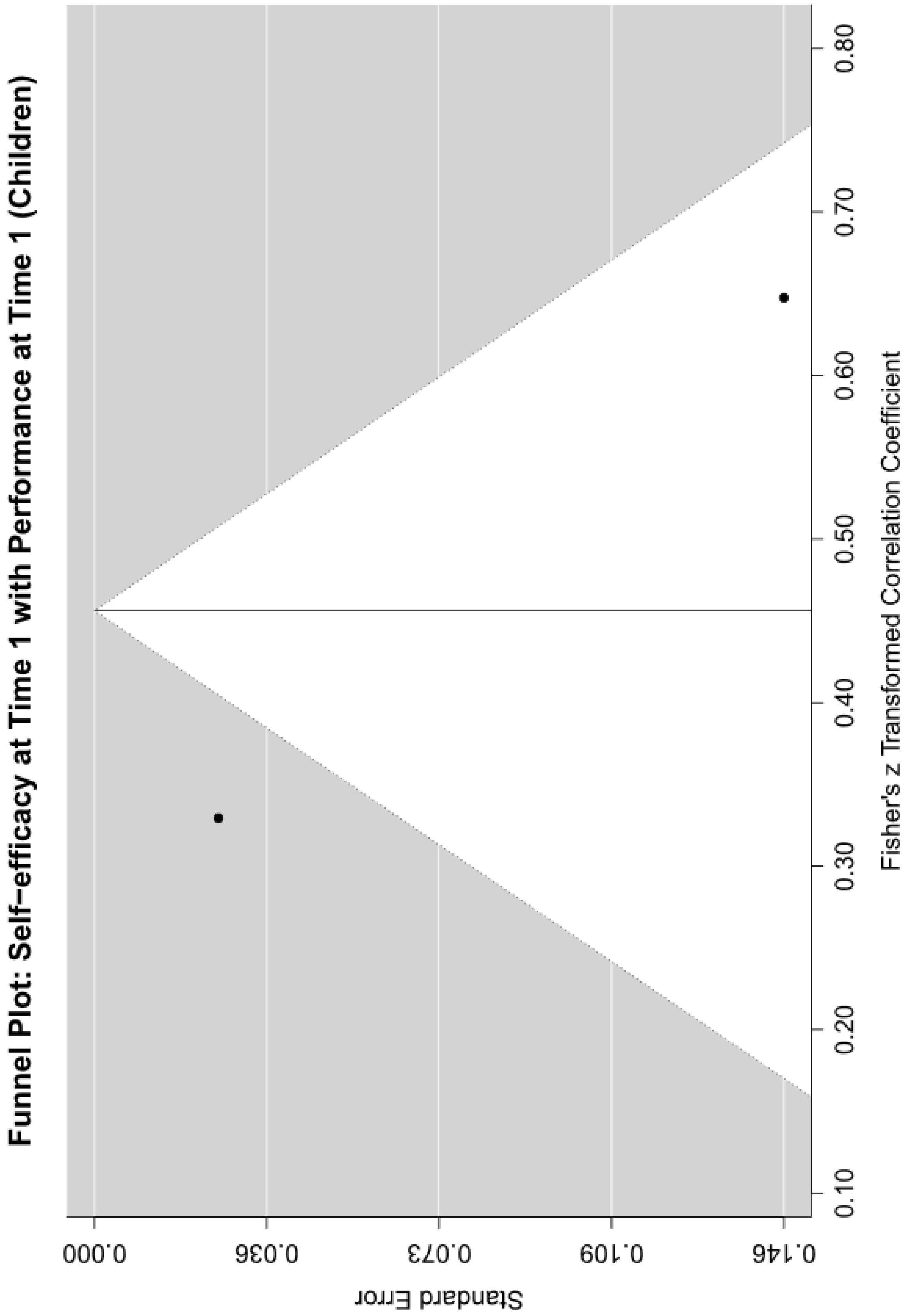

Forest Plot: Self−efficacy at Time 1 with Performance at Time 2 (Adults)
n
Galyon et al., 2012 213 0.03 [−0.11 , 0.16]
Mone et al., 1995 215 0.22 [0.09 , 0.36]
Finney et al., 2003 103 0.23 [0.03 , 0.42]
Daniels et al., 2010 111 0.23 [0.04 , 0.42]
Vancouver et al., 2006 65 0.26 [0.01 , 0.51]
Mone, 1994 252 0.30 [0.17 , 0.42]
Beck et al., 2013 69 0.30 [0.06 , 0.54]
Richard et al., 2006 83 0.35 [0.13 , 0.57]
Meier et al., 1984 71 0.39 [0.15 , 0.63]
RE Model 0.24 [0.16 , 0.32]
−0.20
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient

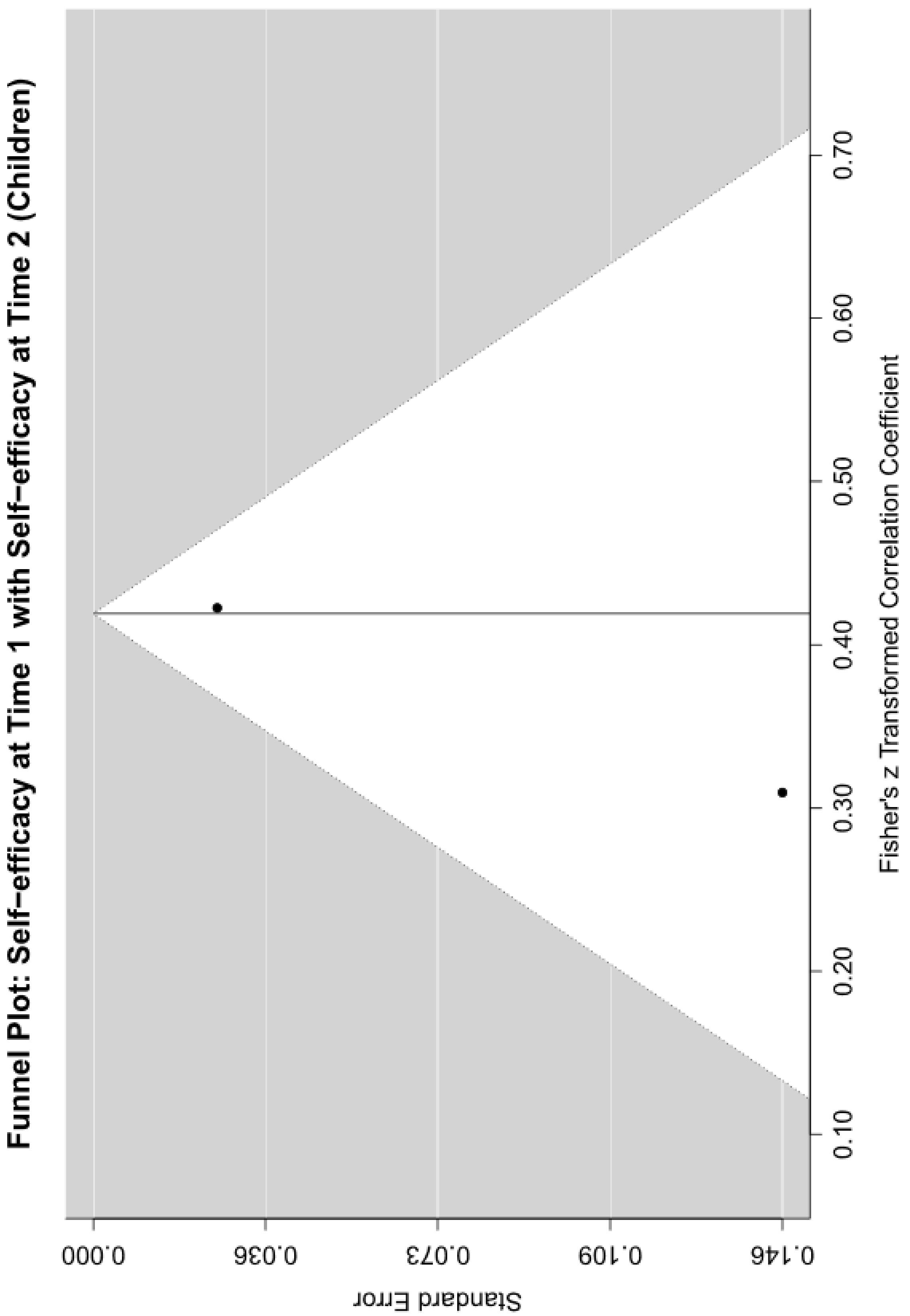

Forest Plot: Self−efficacy at Time 2 with Performance at Time 1 (Adults)
n
Meier et al., 1984
71
0.21 [−0.02 , 0.45]
Galyon et al., 2012
213
0.24 [0.11 , 0.38]
Finney et al., 2003
103
0.43 [0.23 , 0.62]
Vancouver et al., 2006
65
0.43 [0.18 , 0.68]
Daniels et al., 2010
111
0.46 [0.27 , 0.65]
Mone et al., 1995
215
0.48 [0.35 , 0.62]
Mone, 1994
252
0.52 [0.40 , 0.65]
Richard et al., 2006
83
0.52 [0.30 , 0.74]
Beck et al., 2013
69
0.53 [0.29 , 0.77]
RE Model
0.43 [0.35 , 0.51]
−0.20
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient

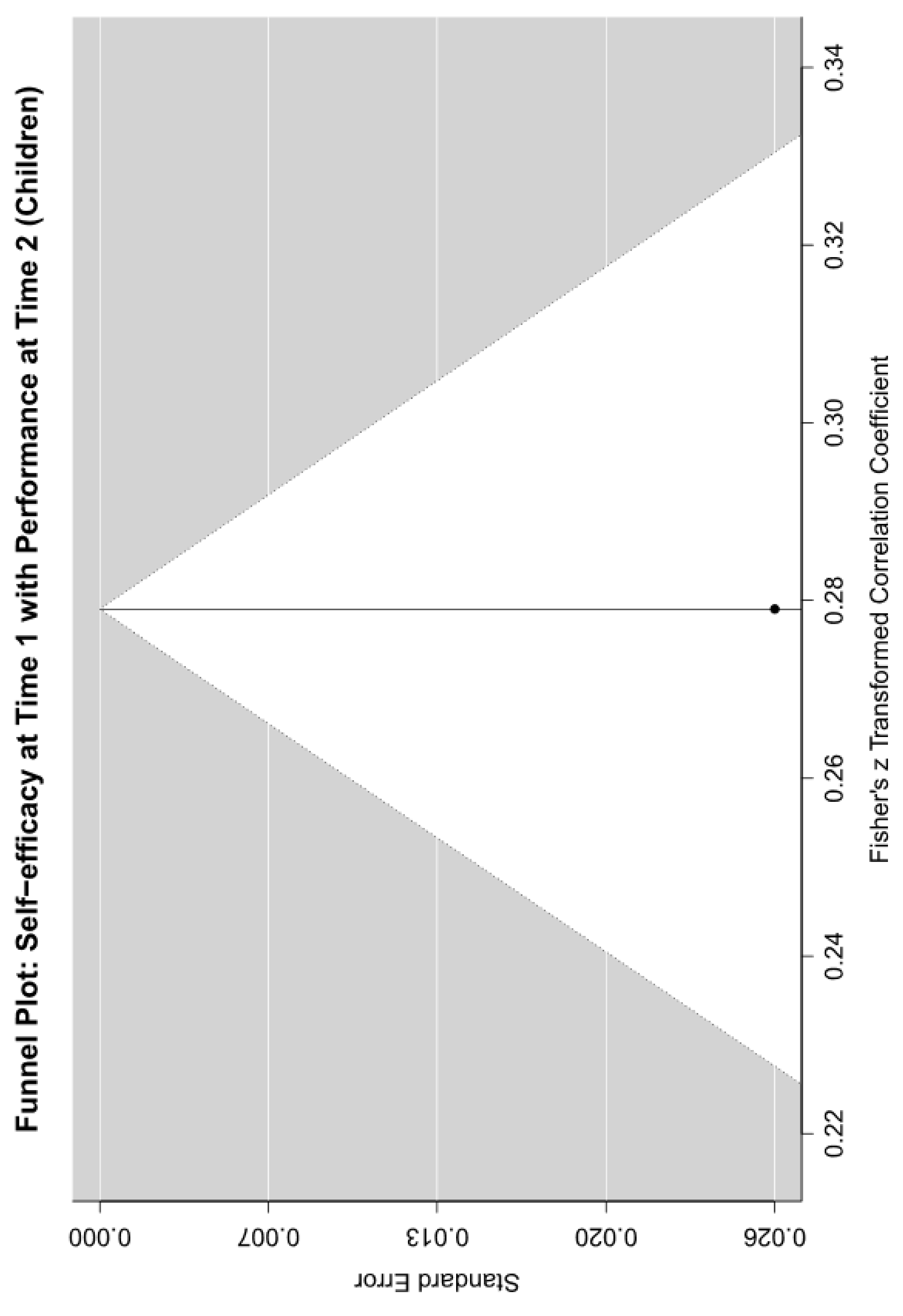


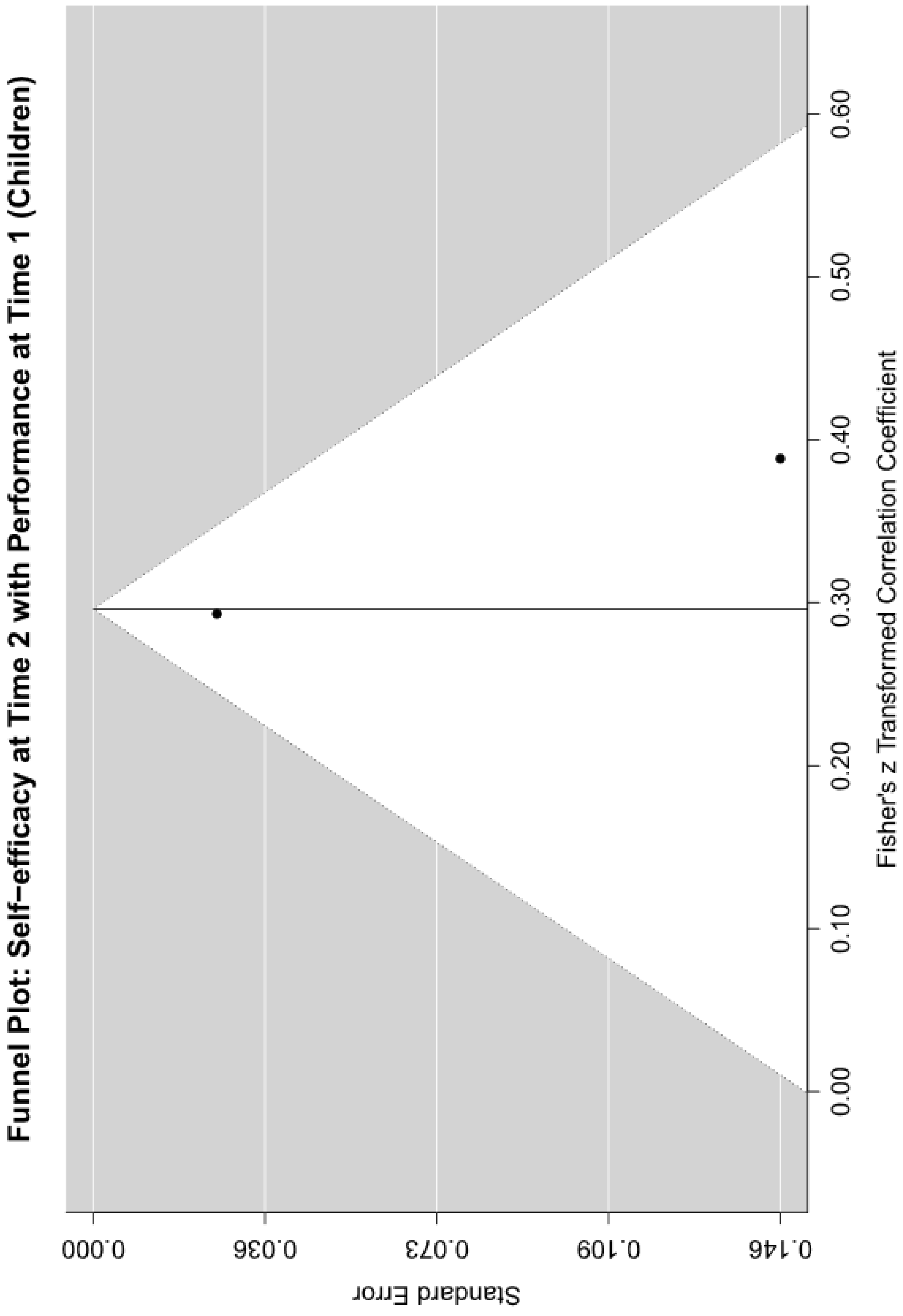
Forest Plot: Performance at Time 1 with Performance at Time 2 (Adults)
n
Vancouver et al., 2006 65 0.19 [−0.06 , 0.44]
Finney et al., 2003 103 0.35 [0.16 , 0.55]
Meier et al., 1984 71 0.50 [0.26 , 0.73]
Mone, 1994 252 0.54 [0.41 , 0.66]
Mone et al., 1995 215 0.59 [0.46 , 0.72]
Daniels et al., 2010 111 0.61 [0.42 , 0.79]
Galyon et al., 2012 213 0.65 [0.51 , 0.78]
Richard et al., 2006 83 0.74 [0.52 , 0.96]
Beck et al., 2013 69 0.86 [0.62 , 1.10]
RE Model 0.56 [0.45 , 0.68]
−0.50
0.00
0.50
1.00
1.50
Fisher's z Transformed Correlation Coefficient

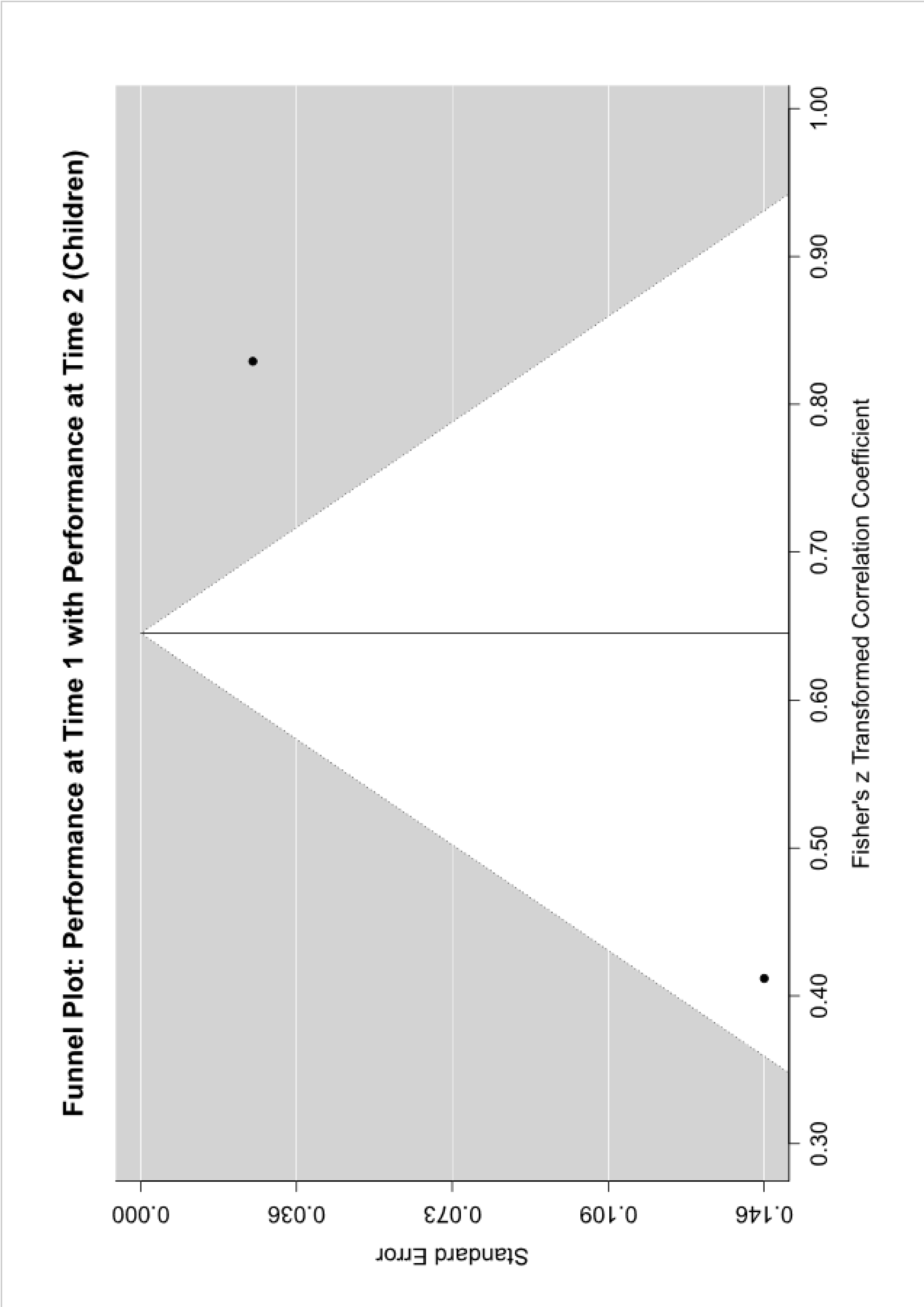

Forest Plot: Self−efficacy at Time 2 with Performance at Time 2 (Adults)
n
Meier et al., 1984
71
0.09 [−0.15 , 0.33]
Galyon et al., 2012
213
0.22 [0.08 , 0.35]
Daniels et al., 2010
111
0.22 [0.03 , 0.41]
Mone et al., 1995
215
0.24 [0.11 , 0.38]
Beck et al., 2013
69
0.32 [0.08 , 0.57]
Vancouver et al., 2006
65
0.33 [0.08 , 0.58]
Mone, 1994
252
0.33 [0.21 , 0.46]
Richard et al., 2006
83
0.41 [0.19 , 0.63]
Finney et al., 2003
103
0.47 [0.28 , 0.67]
RE Model
0.29 [0.23 , 0.35]
−0.20
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient

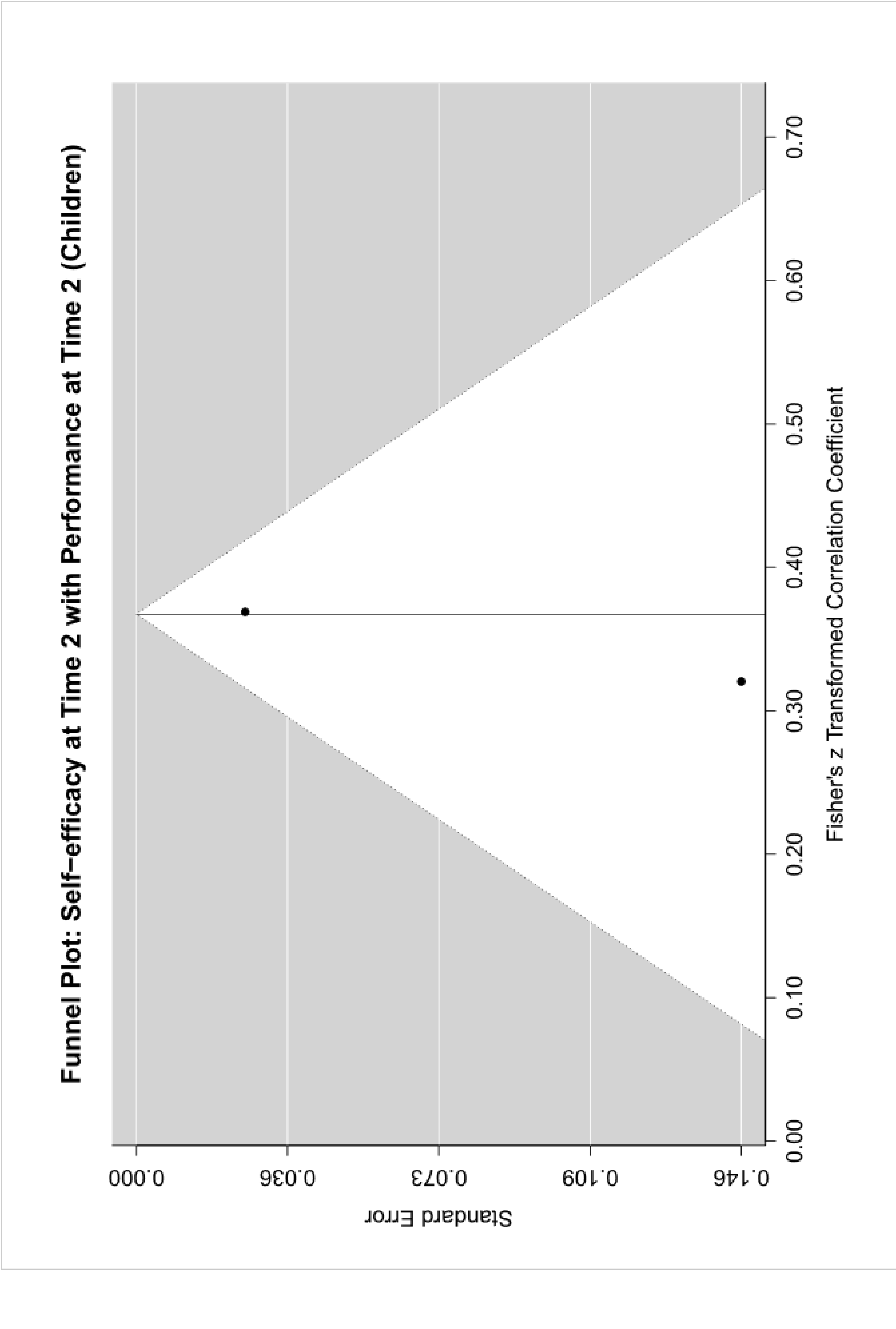

Funnel Plot: Self−efficacy at Time 1 with Performance at Time 1 (Children)
Standard Error
0.000
0.036
0.073
0.109
0.146
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
Fisher's z Transformed Correlation Coefficient

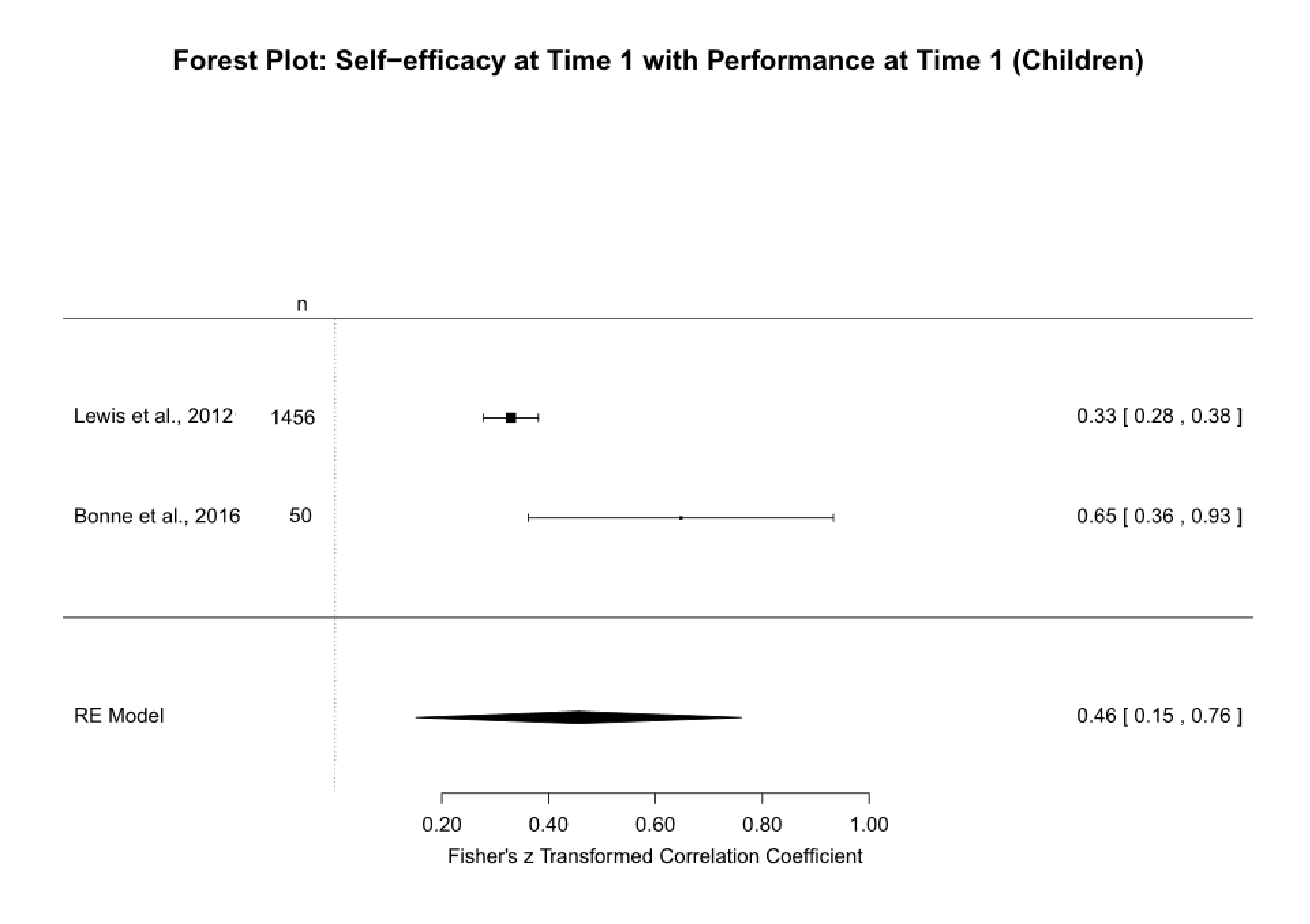

Funnel Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2 (Children)
Standard Error
0.000
0.036
0.073
0.109
0.146
0.10
0.20
0.30
0.40
0.50
0.60
0.70
Fisher's z Transformed Correlation Coefficient

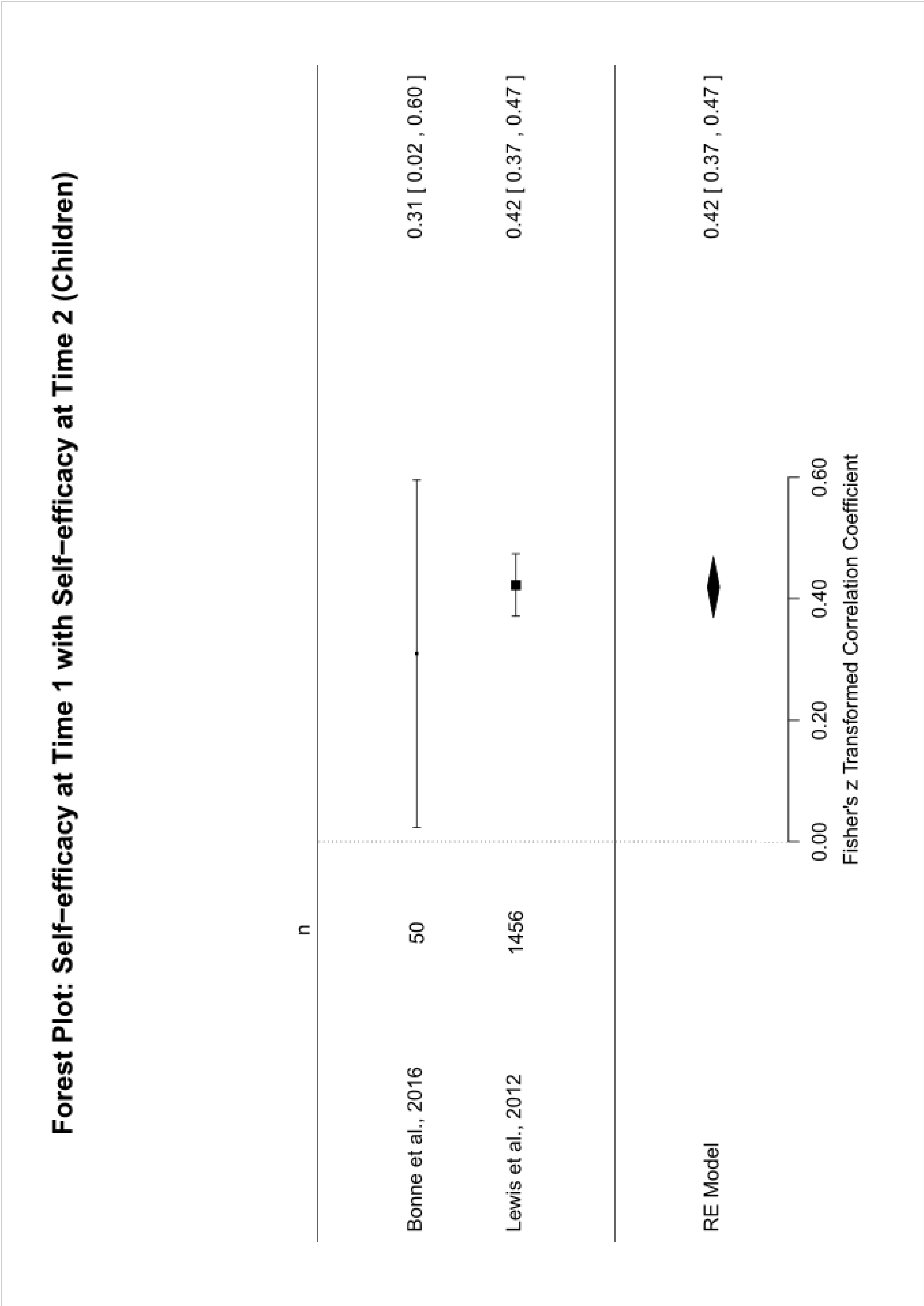

Funnel Plot: Self−efficacy at Time 1 with Performance at Time 2 (Children)
Standard Error
0.000
0.007
0.013
0.020
0.026
Fisher's z Transformed Correlation Coefficient
0.22
0.24
0.26
0.28
0.30
0.32
0.34

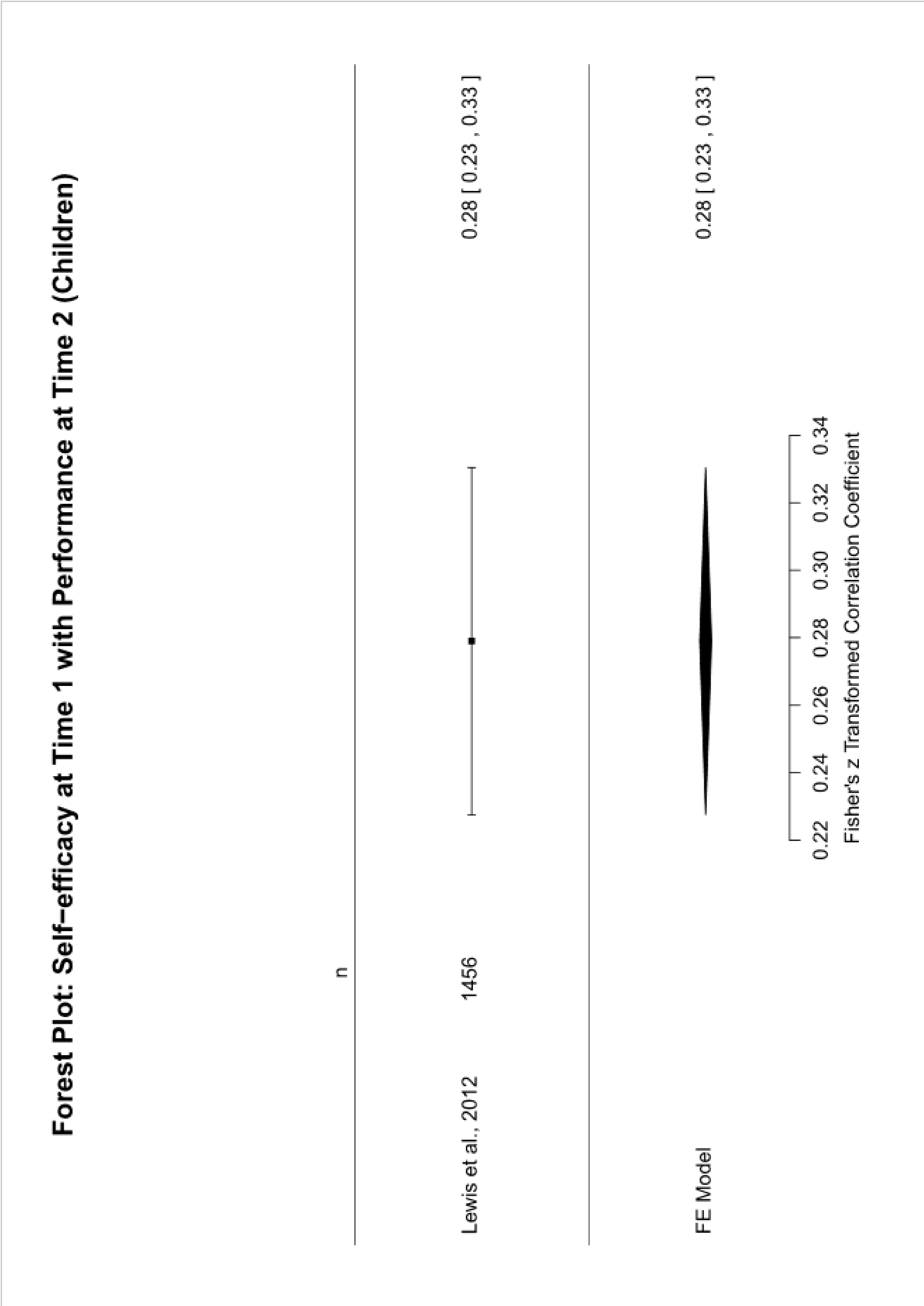

Funnel Plot: Self−efficacy at Time 2 with Performance at Time 1 (Children)
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50
0.60

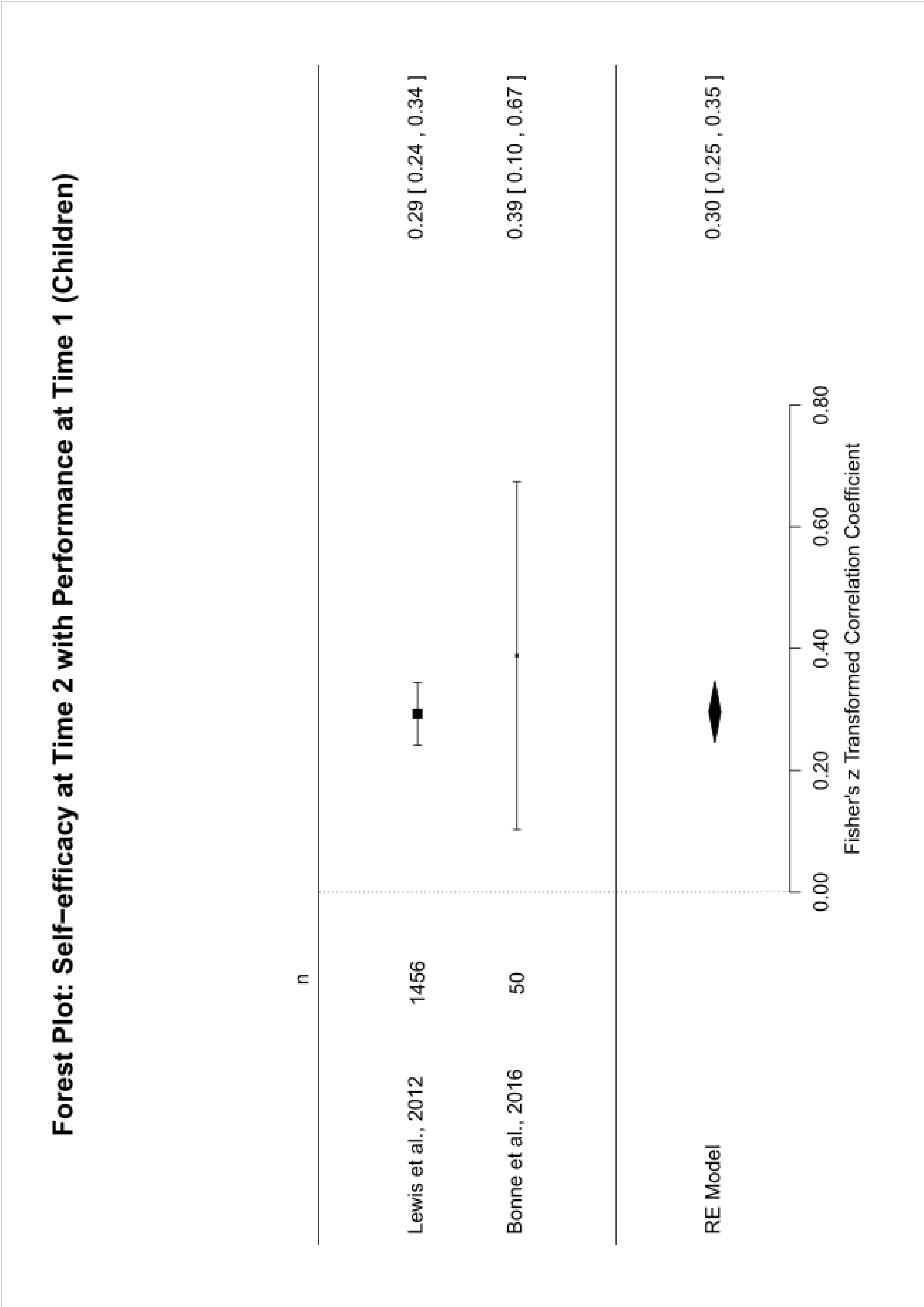

Funnel Plot: Performance at Time 1 with Performance at Time 2 (Children)
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00

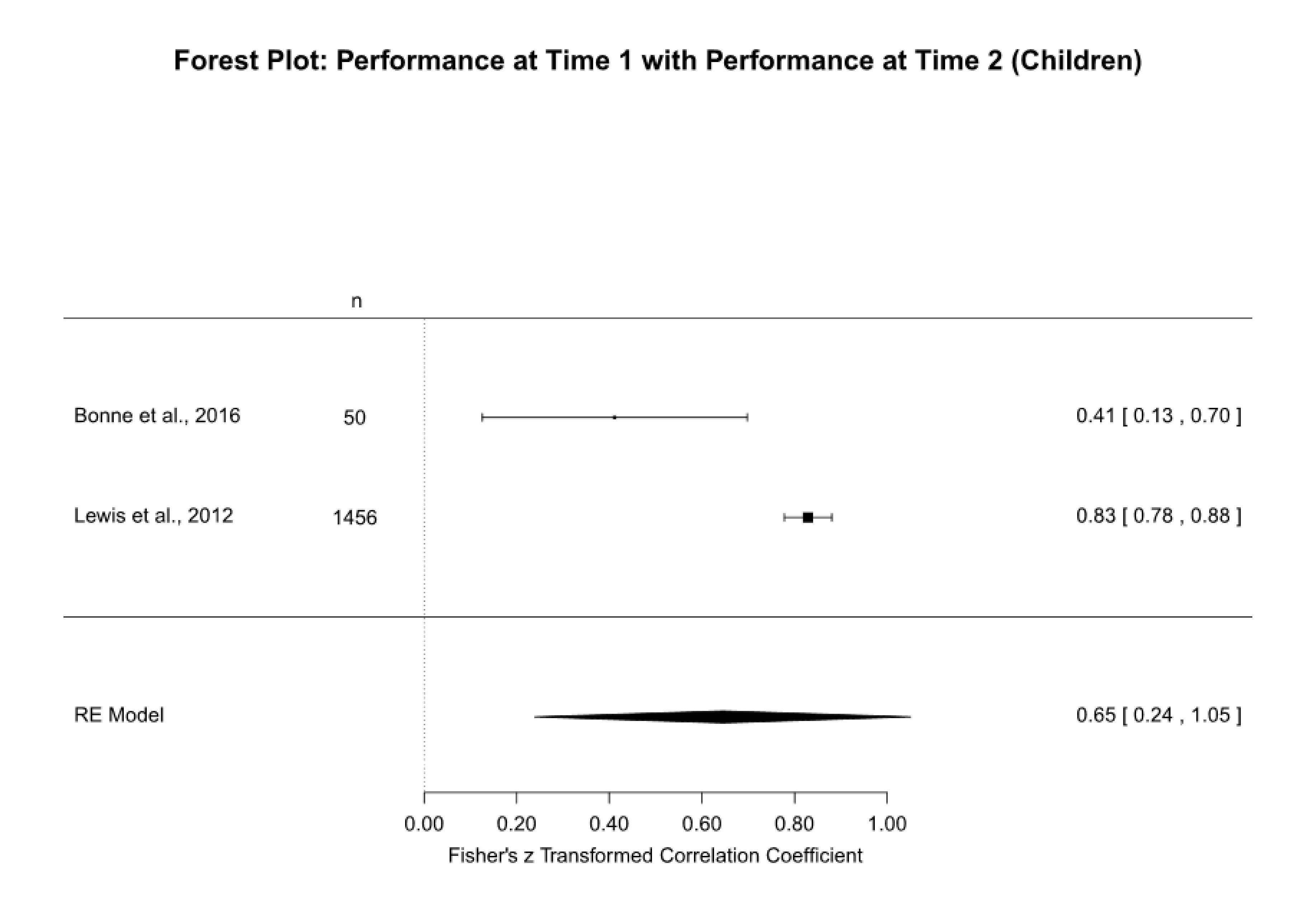

Funnel Plot: Self−efficacy at Time 2 with Performance at Time 2 (Children)
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70

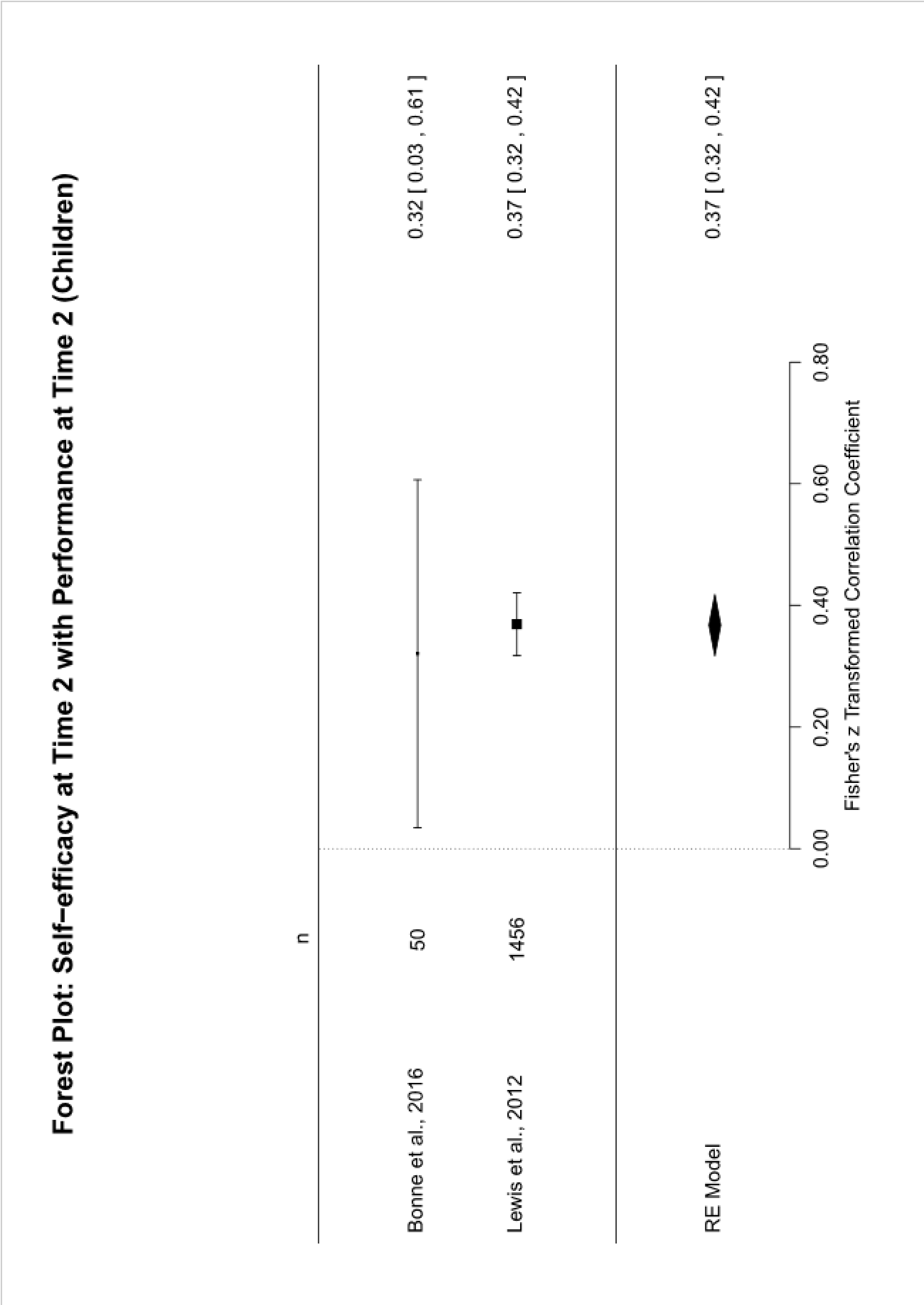

Forest Plot: Self−efficacy at Time 1 with Performance at Time 1 (Children)
n
Lewis et al., 2012
1456
0.33 [0.28 , 0.38]
Bonne et al., 2016
50
0.65 [0.36 , 0.93]
RE Model
0.46 [0.15 , 0.76]
0.20
0.40
0.60
0.80
1.00
Fisher's z Transformed Correlation Coefficient


Forest Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2 (Children)
n
Bonne et al., 2016
50
0.31 [0.02 , 0.60]
Lewis et al., 2012
1456
0.42 [0.37 , 0.47]
RE Model
0.42 [0.37 , 0.47]
0.00
0.20
0.40
0.60
Fisher's z Transformed Correlation Coefficient

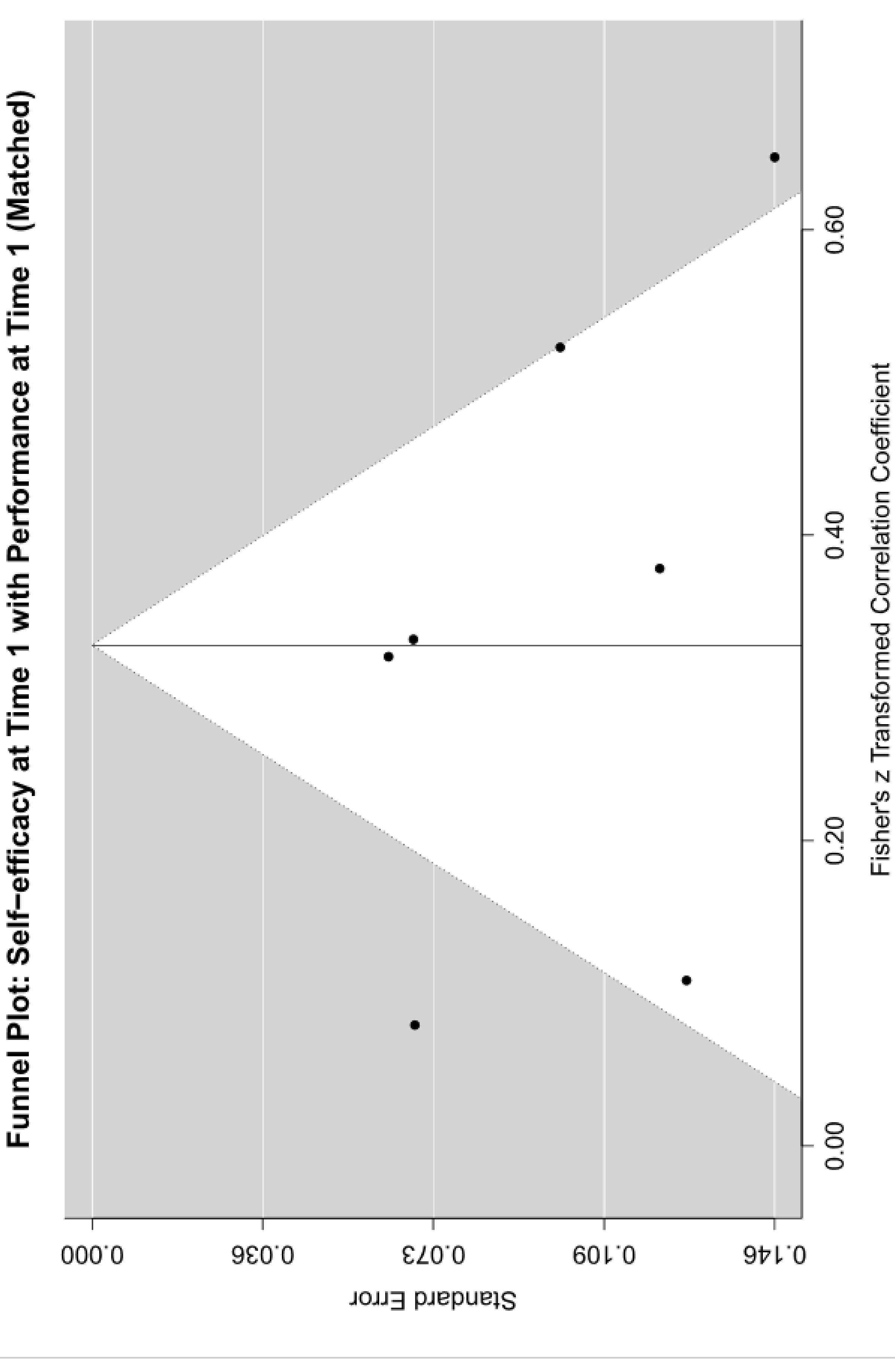

Forest Plot: Self–efficacy at Time 1 with Performance at Time 2 (Children)
n
Lewis et al., 2012
1456
0.28 [0.23 , 0.33]
FE Model
0.28 [0.23 , 0.33]
0.22
0.24
0.26
0.28
0.30
0.32
0.34
Fisher's z Transformed Correlation Coefficient


Forest Plot: Self–efficacy at Time 2 with Performance at Time 1 (Children)
n
Lewis et al., 2012
1456
0.29 [0.24 , 0.34]
Bonne et al., 2016
50
0.39 [0.10 , 0.67]
RE Model
0.30 [0.25 , 0.35]
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient

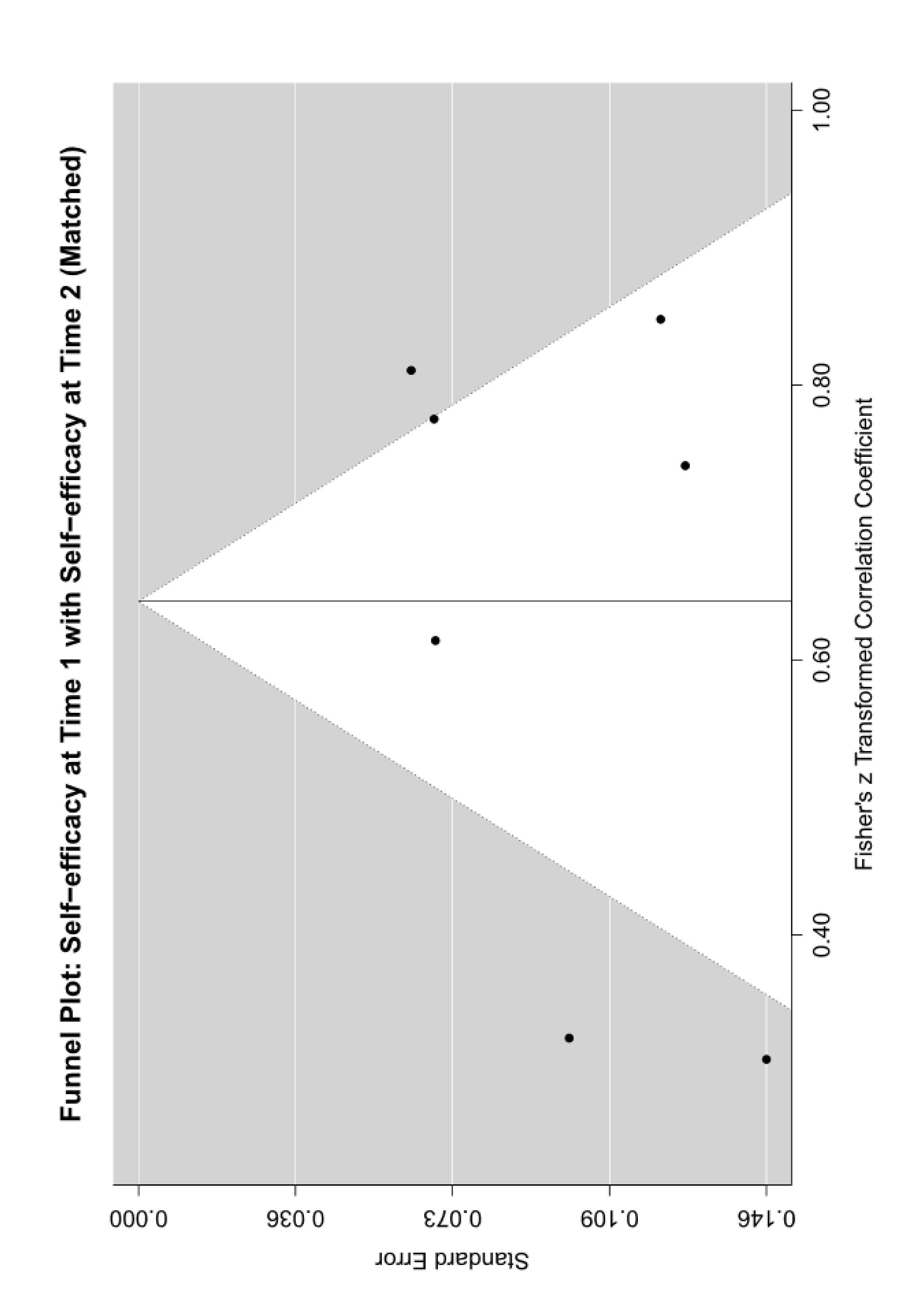

Forest Plot: Performance at Time 1 with Performance at Time 2 (Children)
n
Bonne et al., 2016
50
0.41 [0.13 , 0.70]
Lewis et al., 2012
1456
0.83 [0.78 , 0.88]
RE Model
0.65 [0.24 , 1.05]
0.00
0.20
0.40
0.60
0.80
1.00
Fisher's z Transformed Correlation Coefficient



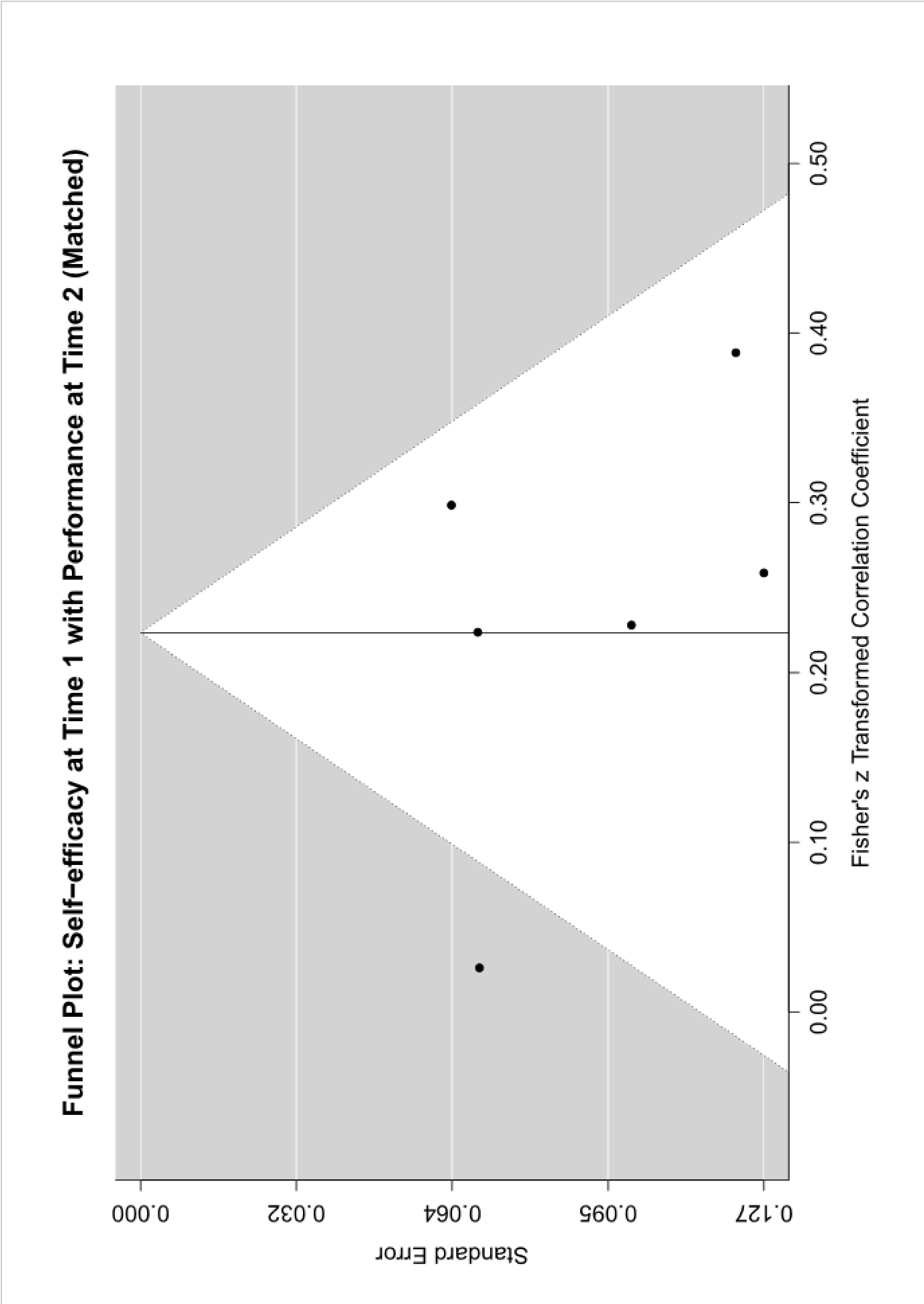
Forest Plot: Self–efficacy at Time 2 with Performance at Time 2 (Children)
n
Bonne et al., 2016
50
0.32 [0.03 , 0.61]
Lewis et al., 2012
1456
0.37 [0.32 , 0.42]
RE Model
0.37 [0.32 , 0.42]
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient

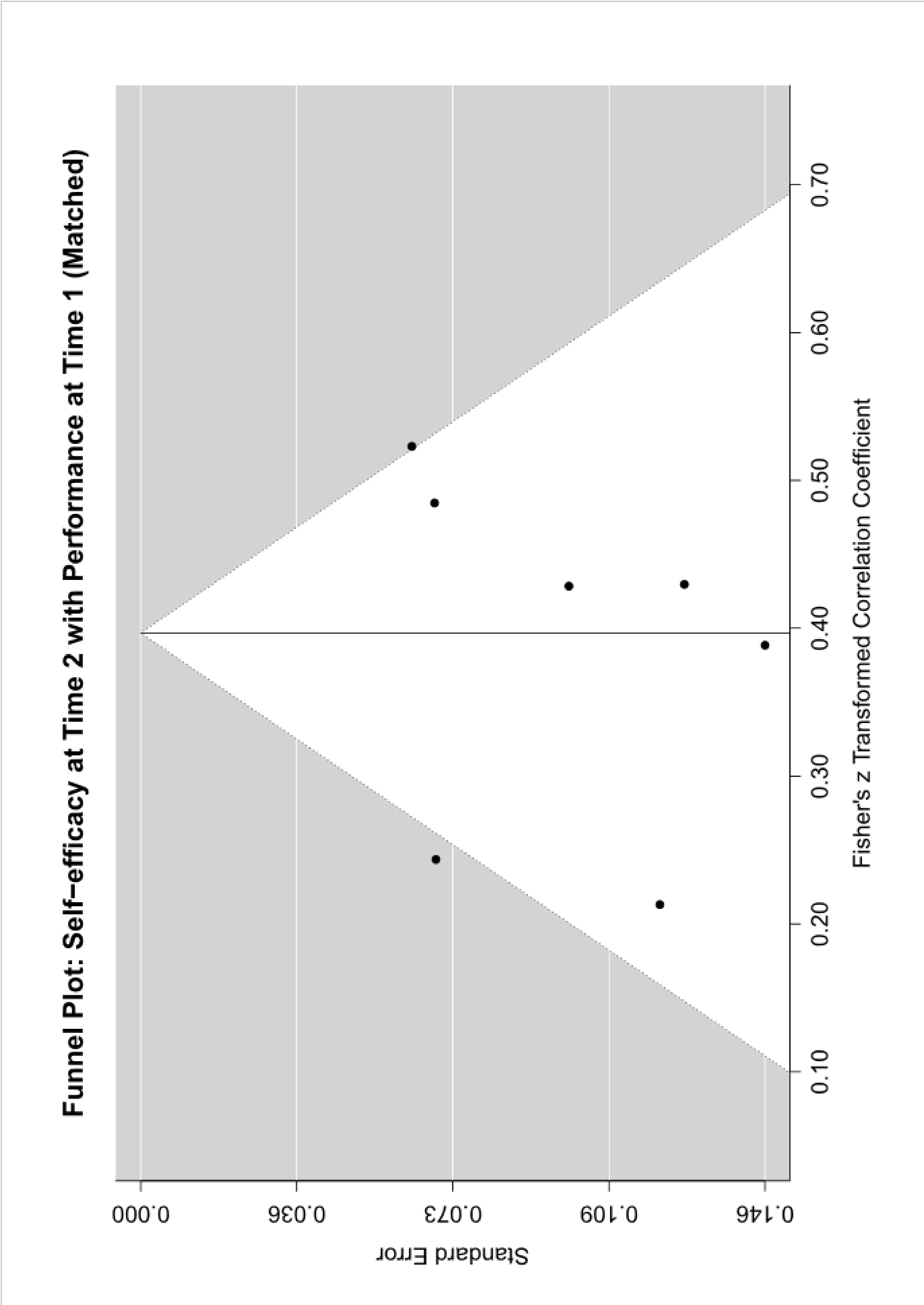


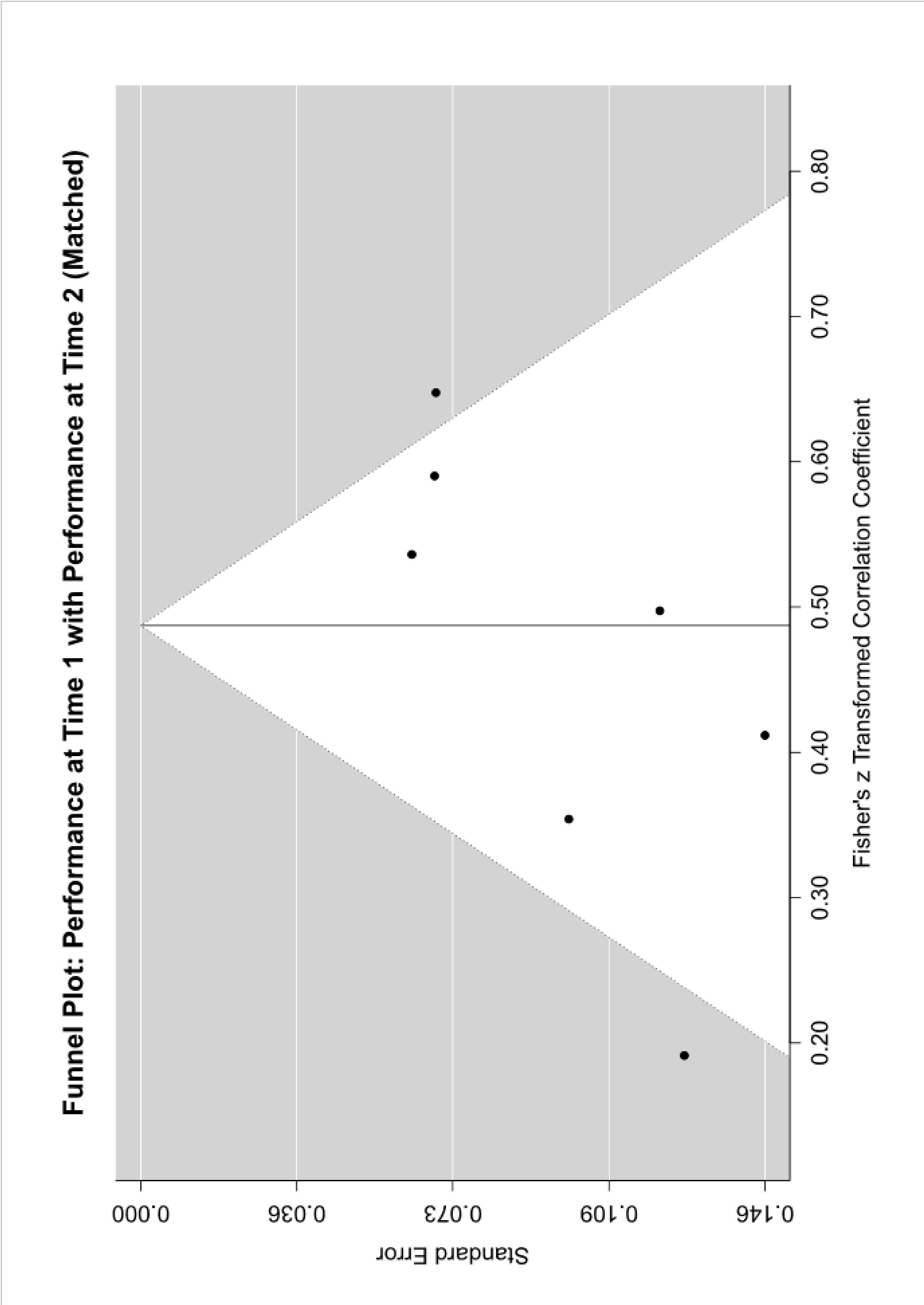
Funnel Plot: Self−efficacy at Time 1 with Performance at Time 1 (Matched)
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.00
0.20
0.40
0.60

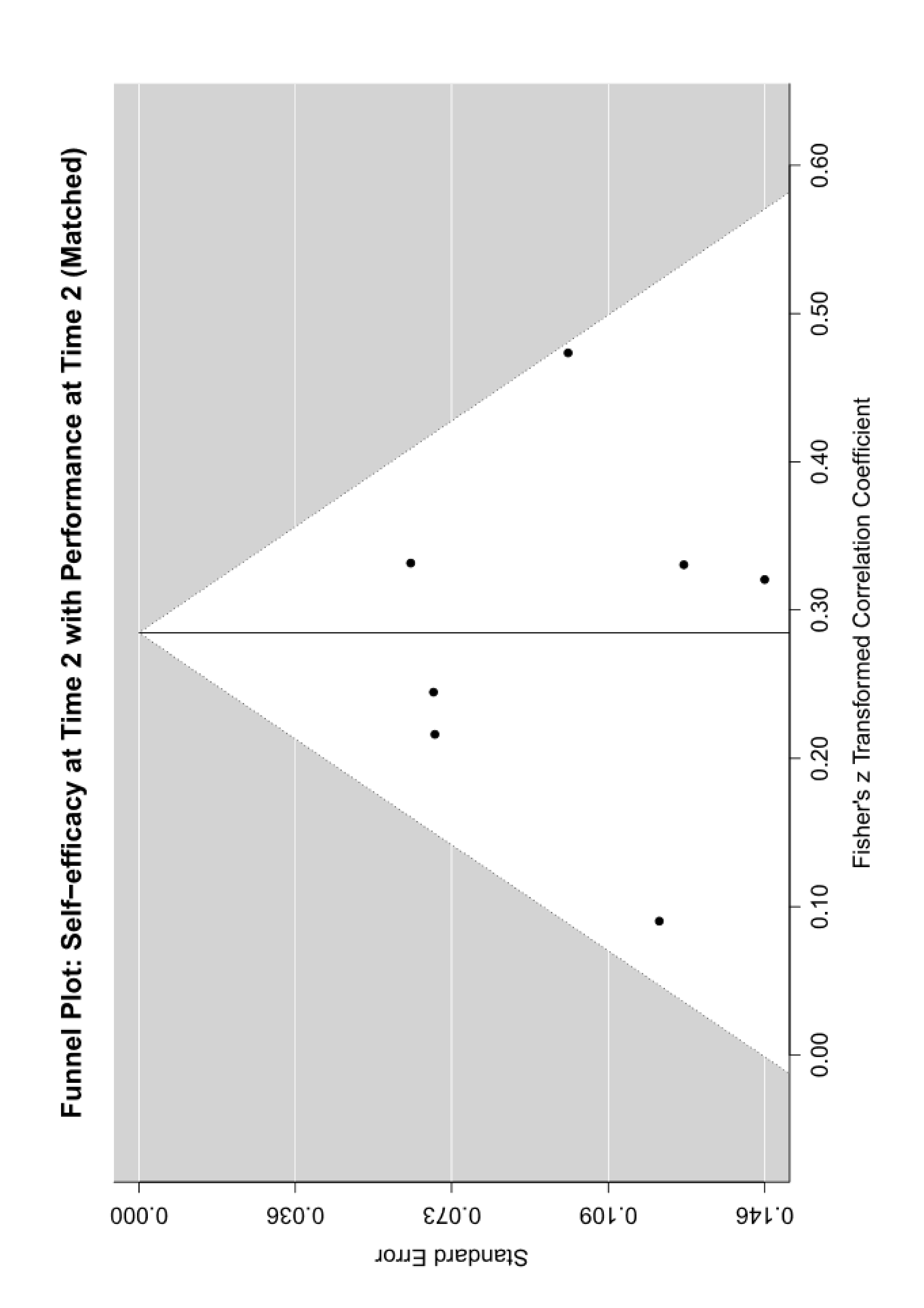


Funnel Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2 (Matched)
Standard Error
0.000
0.036
0.073
0.109
0.146
0.40
0.60
0.80
1.00
Fisher's z Transformed Correlation Coefficient

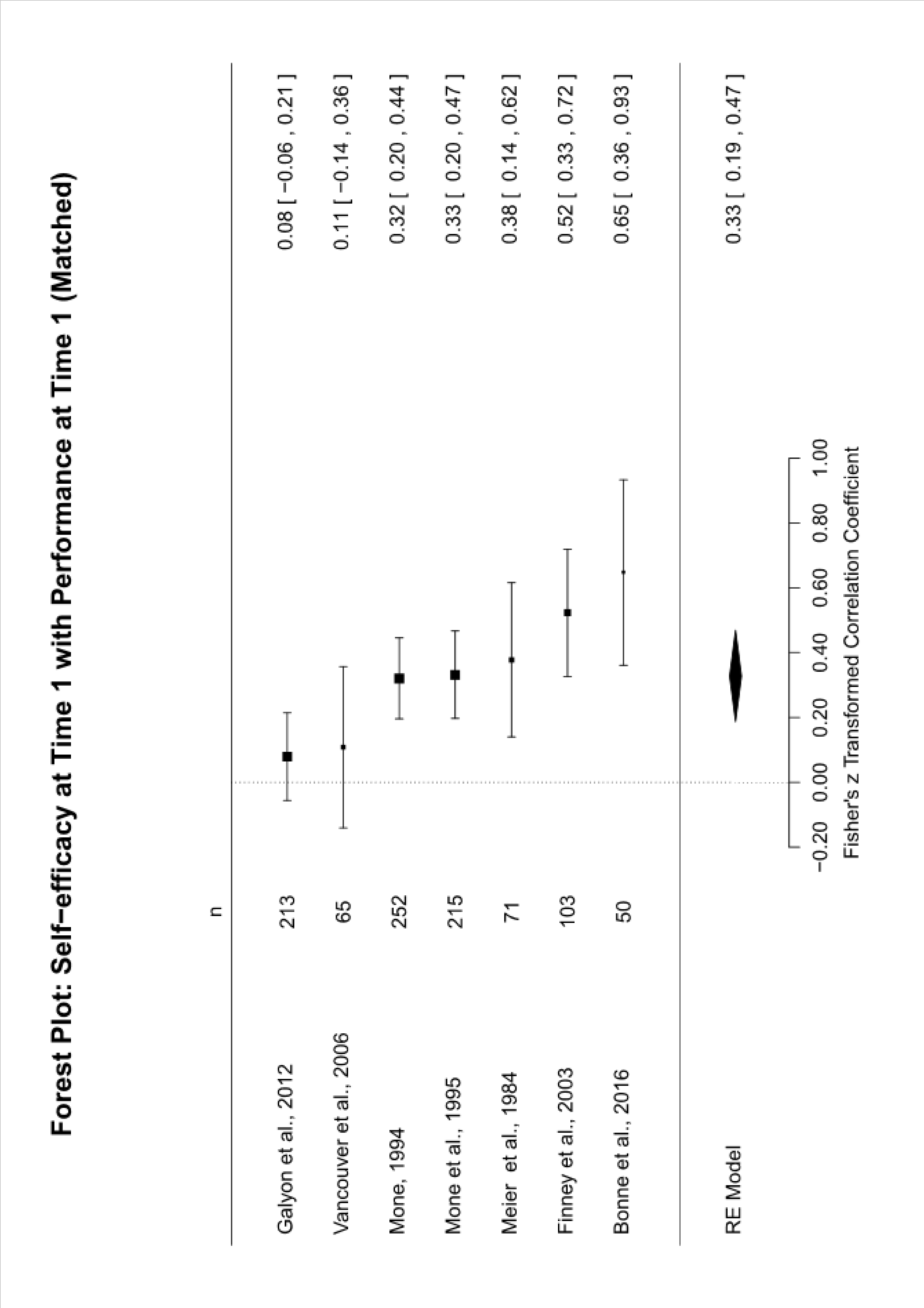

Funnel Plot: Self−efficacy at Time 1 with Performance at Time 2 (Matched)
Standard Error
0.000
0.032
0.064
0.095
0.127
0.00
0.10
0.20
0.30
0.40
0.50
Fisher's z Transformed Correlation Coefficient


Funnel Plot: Self−efficacy at Time 2 with Performance at Time 1 (Matched)
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.10
0.20
0.30
0.40
0.50
0.60
0.70

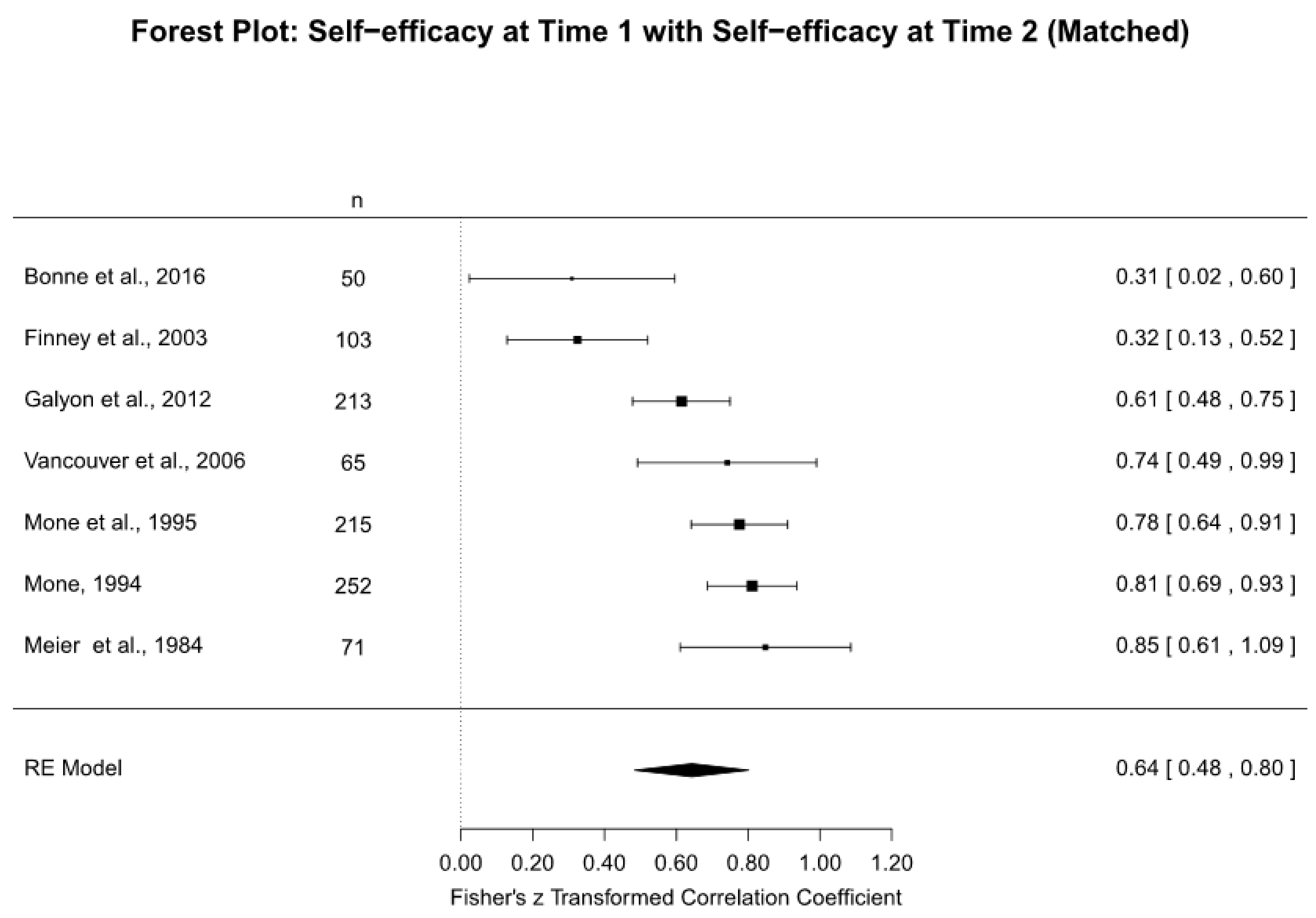

Funnel Plot: Performance at Time 1 with Performance at Time 2 (Matched)
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.20
0.30
0.40
0.50
0.60
0.70
0.80



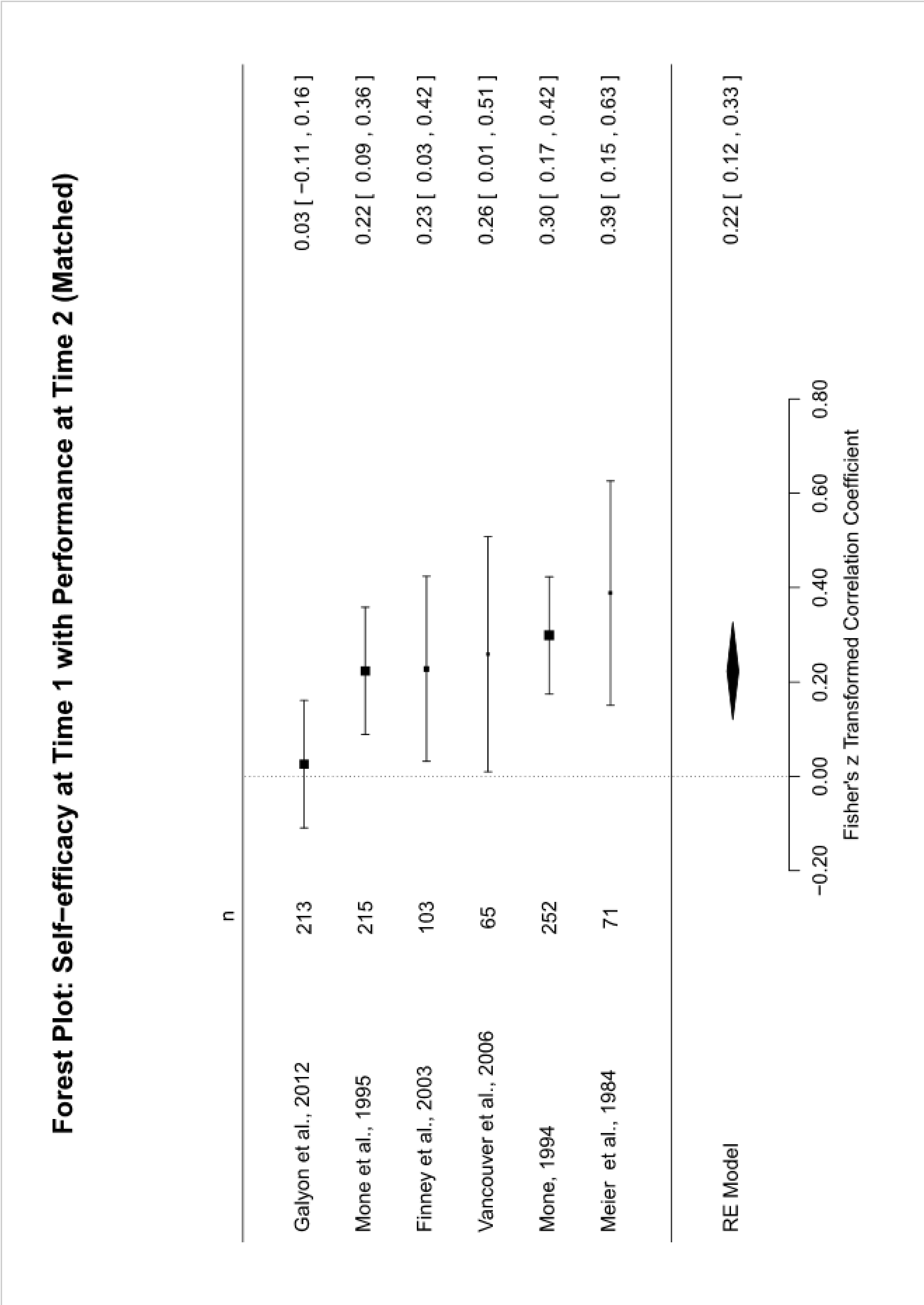
Funnel Plot: Self−efficacy at Time 2 with Performance at Time 2 (Matched)
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50
0.60



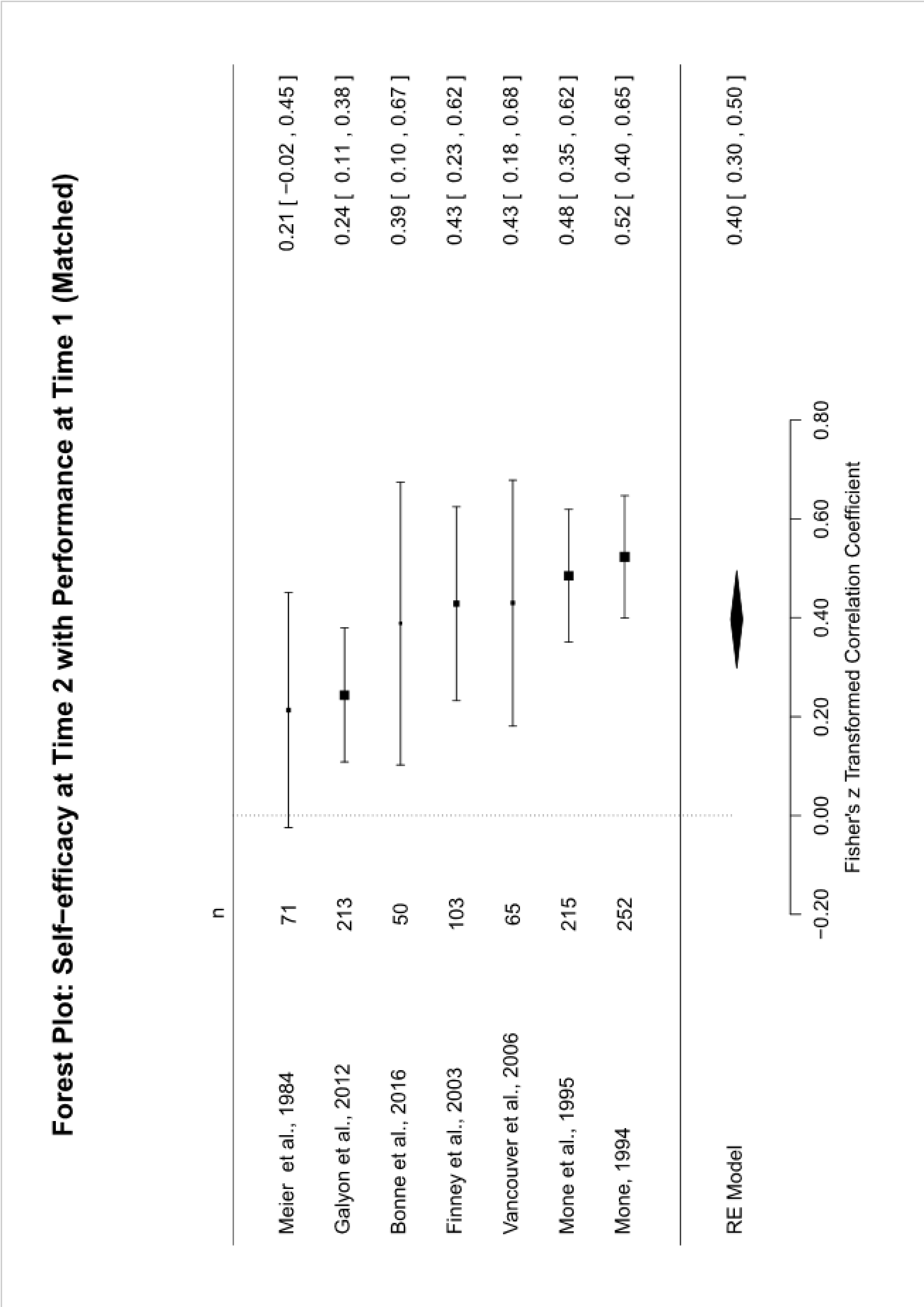
Forest Plot: Self−efficacy at Time 1 with Performance at Time 1 (Matched)
n
Galyon et al., 2012
213
0.08 [−0.06 , 0.21]
Vancouver et al., 2006
65
0.11 [−0.14 , 0.36]
Mone, 1994
252
0.32 [0.20 , 0.44]
Mone et al., 1995
215
0.33 [0.20 , 0.47]
Meier et al., 1984
71
0.38 [0.14 , 0.62]
Finney et al., 2003
103
0.52 [0.33 , 0.72]
Bonne et al., 2016
50
0.65 [0.36 , 0.93]
RE Model
0.33 [0.19 , 0.47]
−0.20 0.00 0.20 0.40 0.60 0.80 1.00
Fisher's z Transformed Correlation Coefficient



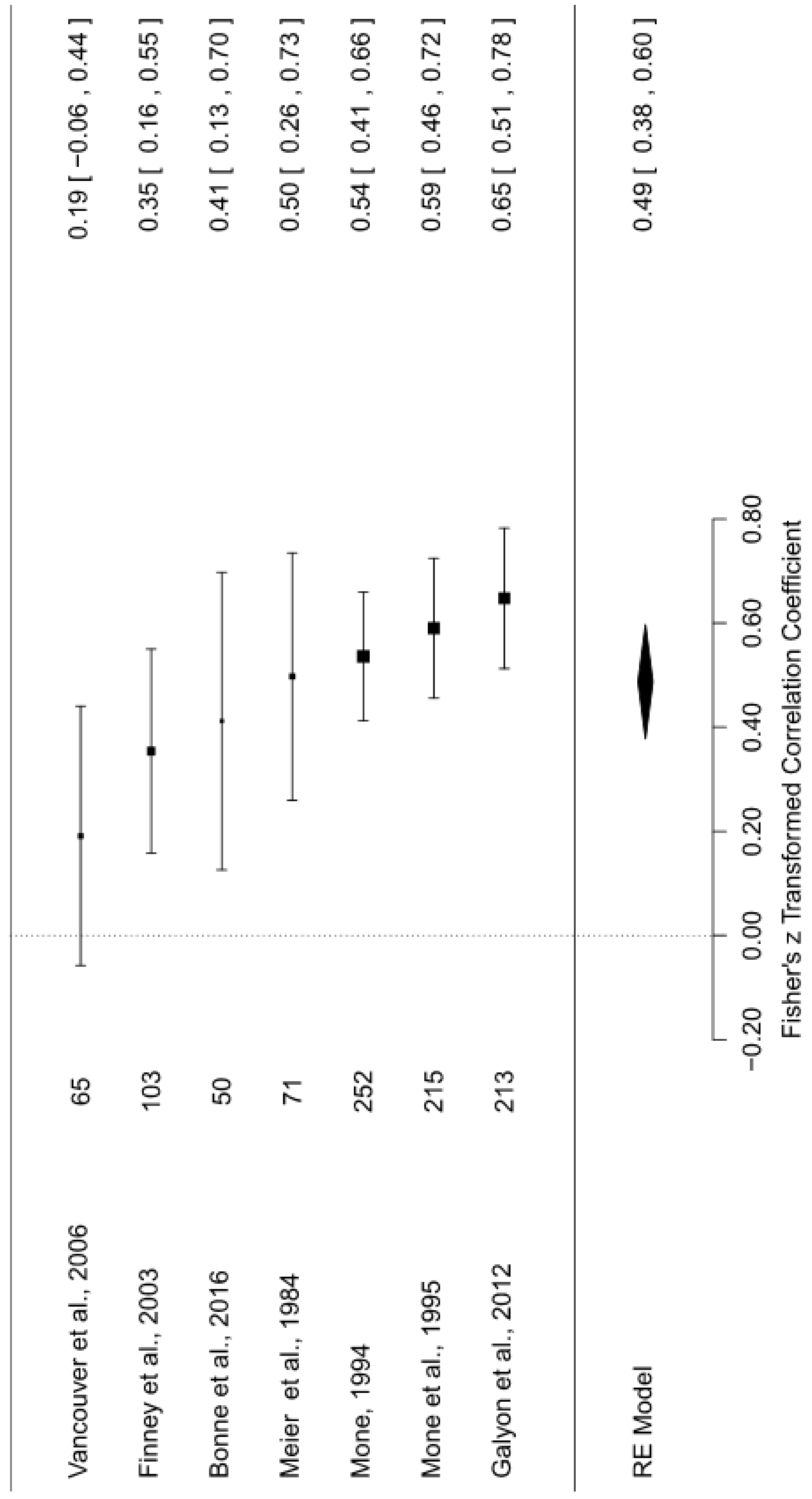
Forest Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2 (Matched)
n
Bonne et al., 2016 50 0.31 [0.02 , 0.60]
Finney et al., 2003 103 0.32 [0.13 , 0.52]
Galyon et al., 2012 213 0.61 [0.48 , 0.75]
Vancouver et al., 2006 65 0.74 [0.49 , 0.99]
Mone et al., 1995 215 0.78 [0.64 , 0.91]
Mone, 1994 252 0.81 [0.69 , 0.93]
Meier et al., 1984 71 0.85 [0.61 , 1.09]
RE Model 0.64 [0.48 , 0.80]
0.00 0.20 0.40 0.60 0.80 1.00 1.20
Fisher's z Transformed Correlation Coefficient



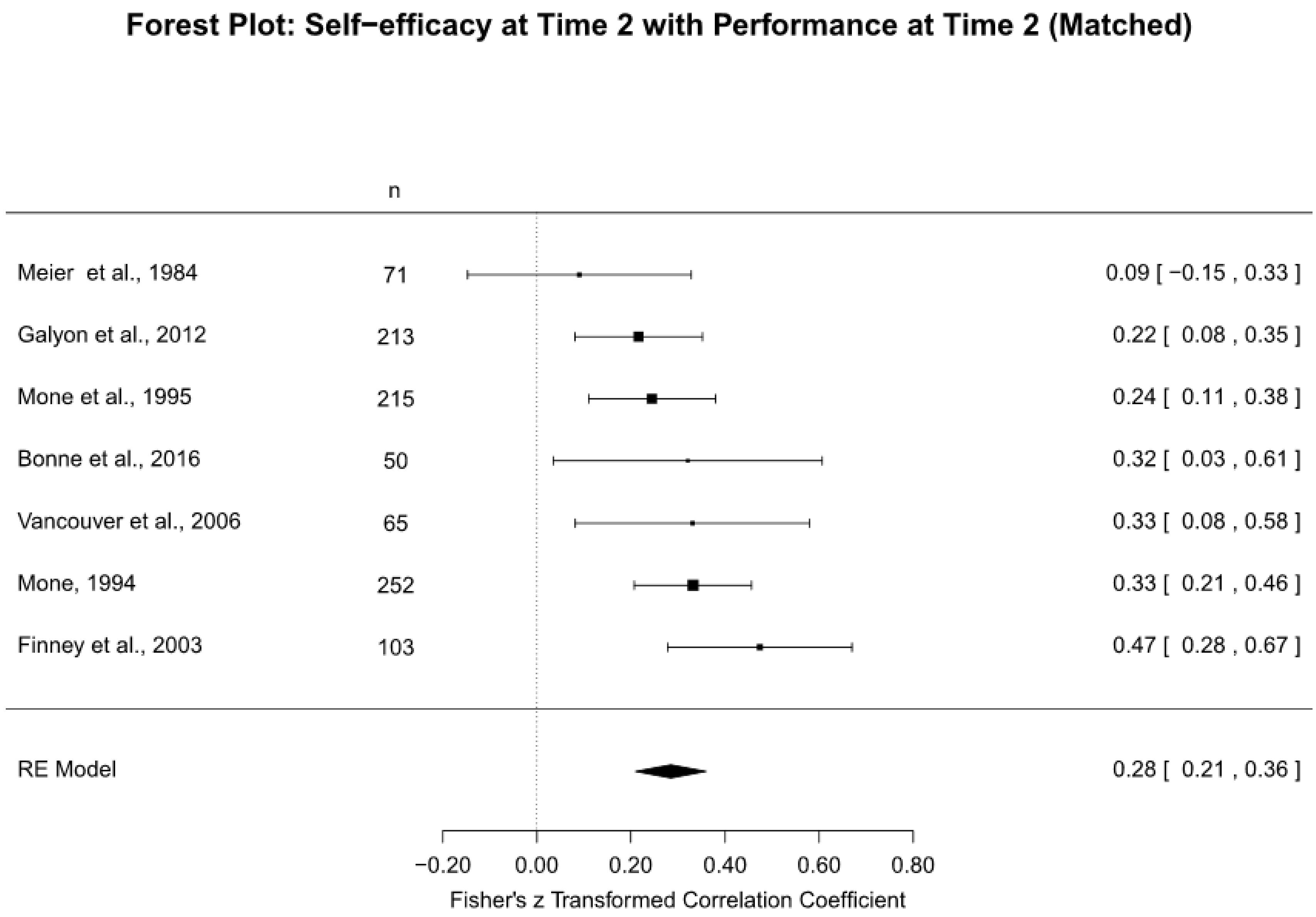
Forest Plot: Self–efficacy at Time 1 with Performance at Time 2 (Matched)
n
Galyon et al., 2012
213
0.03 [−0.11 , 0.16]
Mone et al., 1995
215
0.22 [0.09 , 0.36]
Finney et al., 2003
103
0.23 [0.03 , 0.42]
Vancouver et al., 2006
65
0.26 [0.01 , 0.51]
Mone, 1994
252
0.30 [0.17 , 0.42]
Meier et al., 1984
71
0.39 [0.15 , 0.63]
RE Model
0.22 [0.12 , 0.33]
−0.20
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient

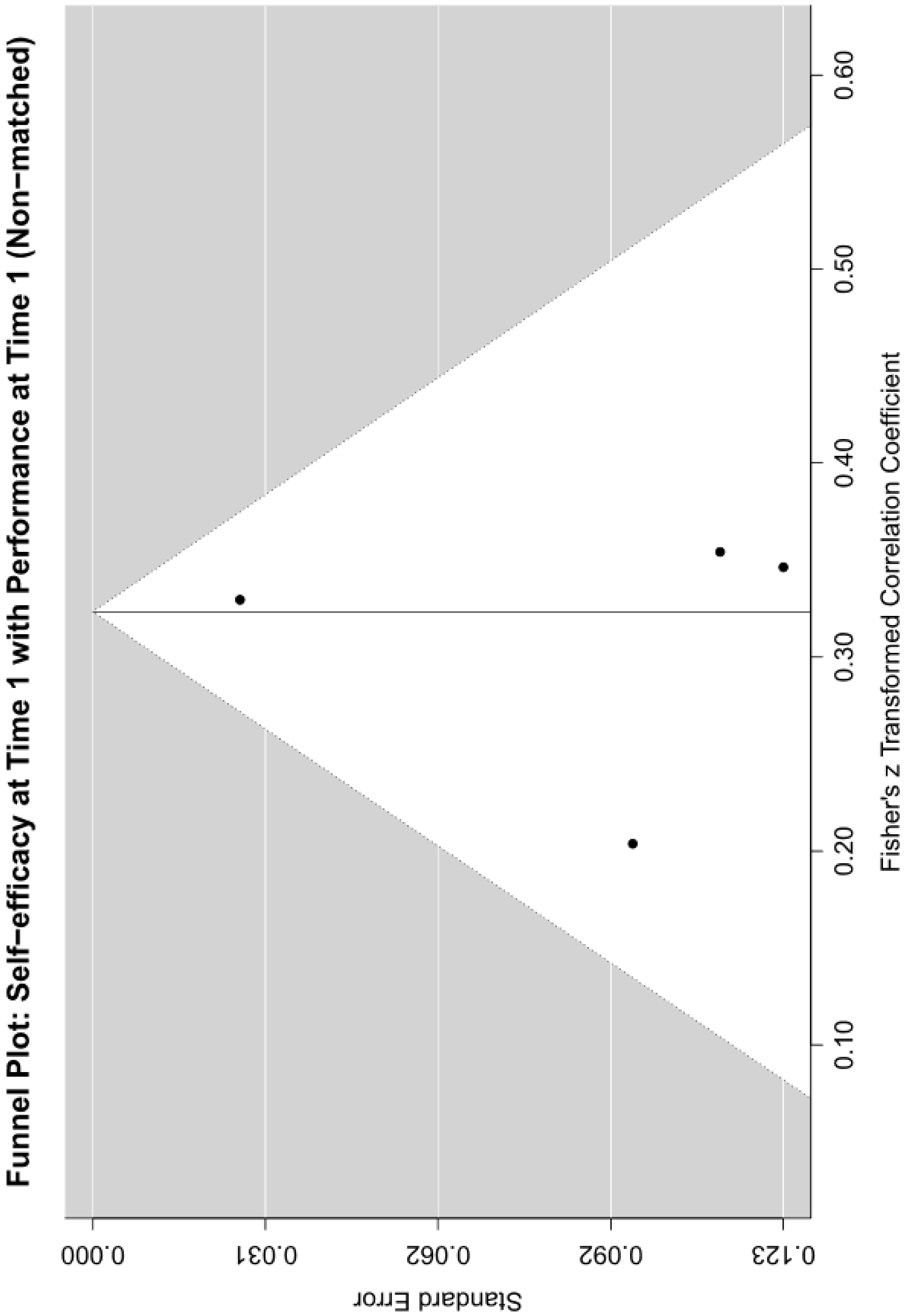


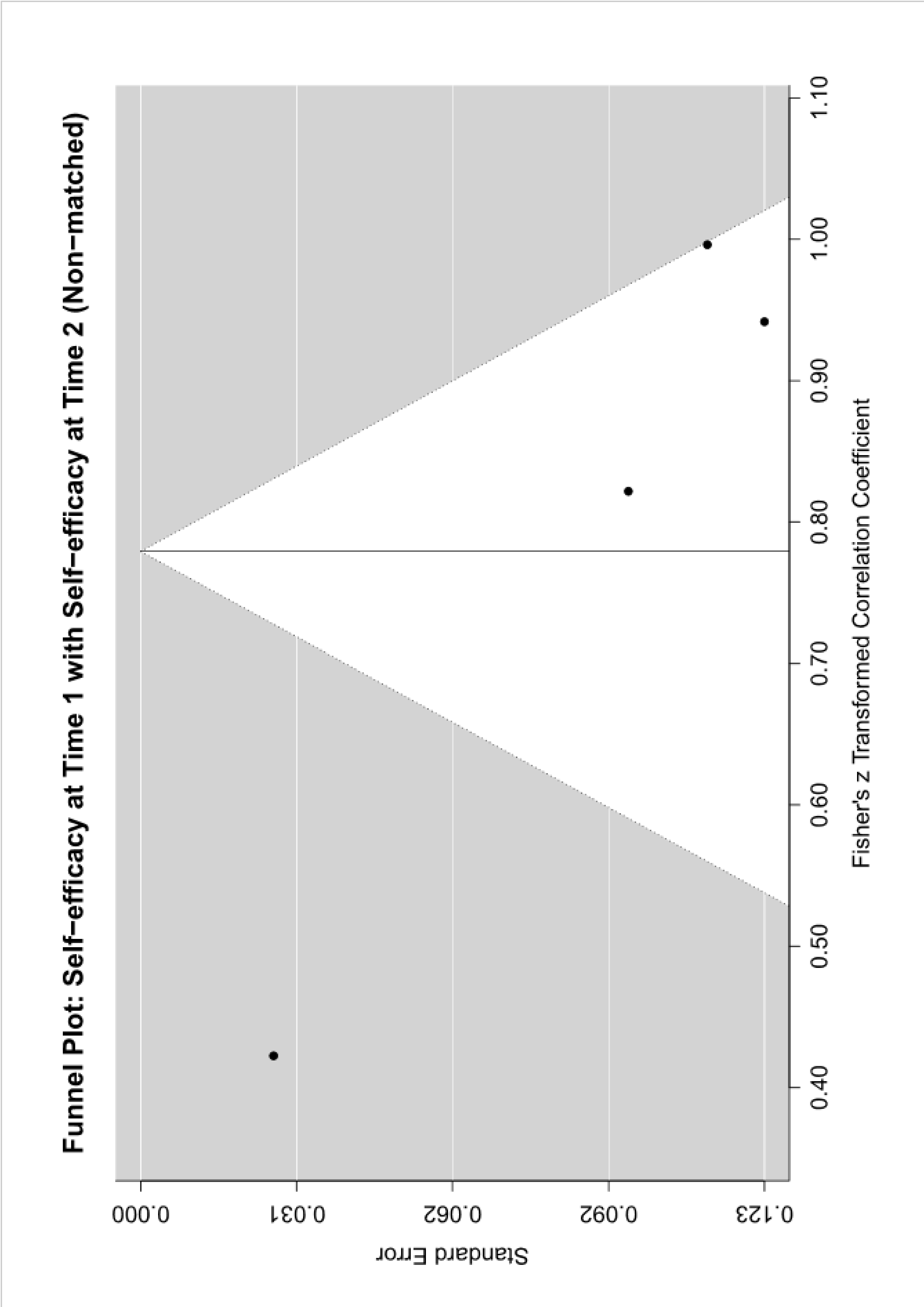
Forest Plot: Self−efficacy at Time 2 with Performance at Time 1 (Matched)
n
Meier et al., 1984
71
0.21 [−0.02 , 0.45]
Galyon et al., 2012
213
0.24 [0.11 , 0.38]
Bonne et al., 2016
50
0.39 [0.10 , 0.67]
Finney et al., 2003
103
0.43 [0.23 , 0.62]
Vancouver et al., 2006
65
0.43 [0.18 , 0.68]
Mone et al., 1995
215
0.48 [0.35 , 0.62]
Mone, 1994
252
0.52 [0.40 , 0.65]
RE Model
0.40 [0.30 , 0.50]
−0.20
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient

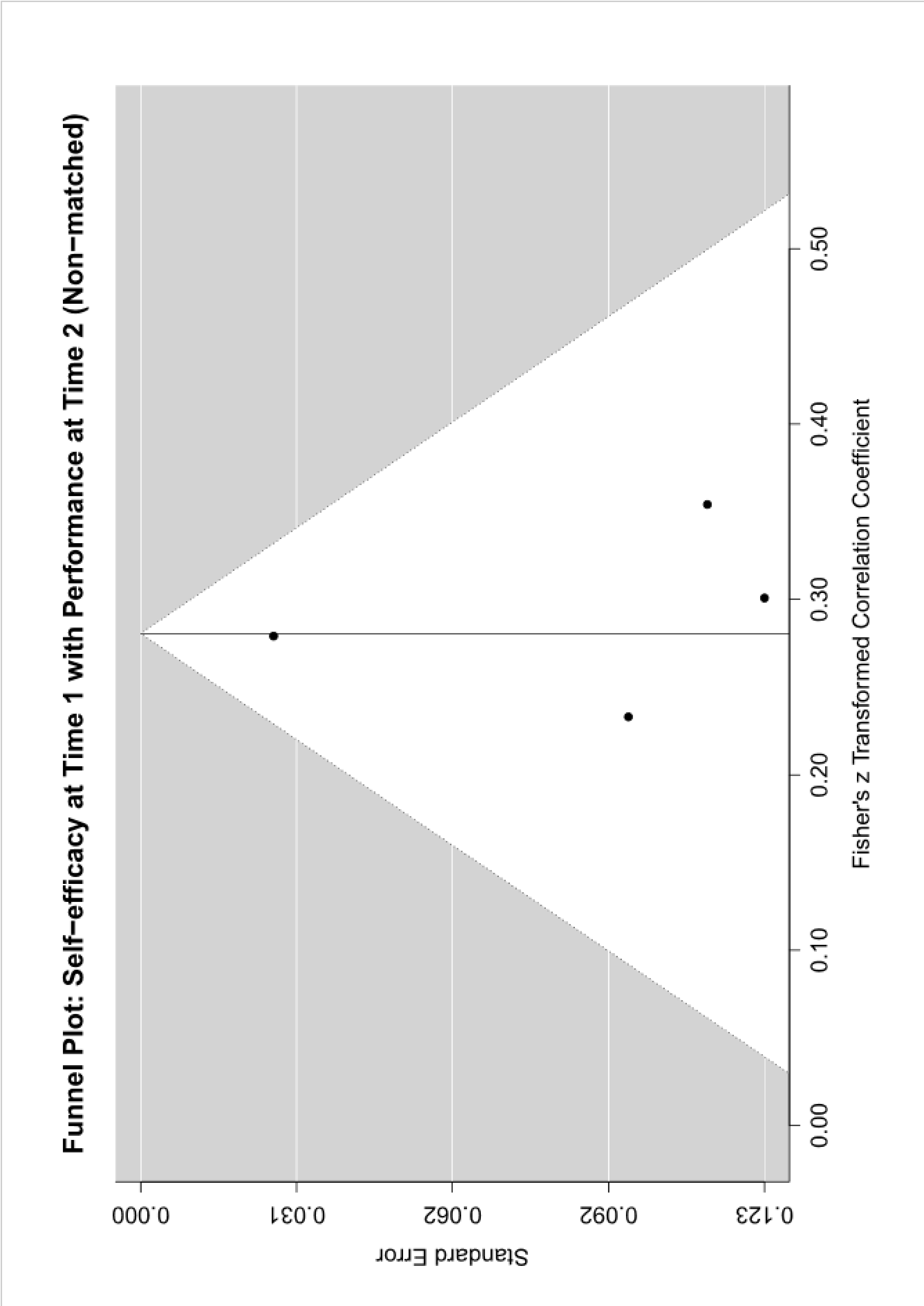


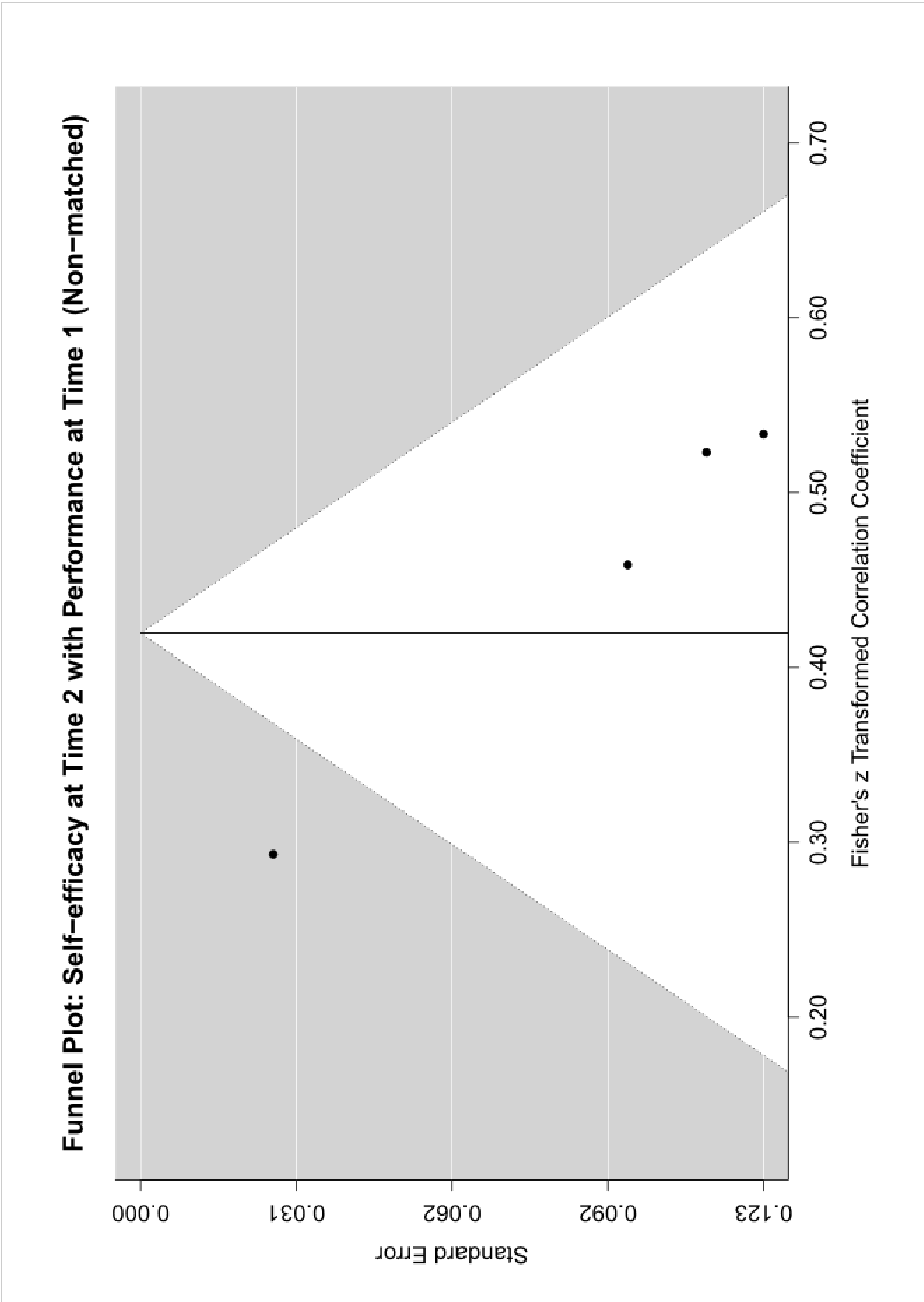
Forest Plot: Performance at Time 1 with Performance at Time 2 (Matched)

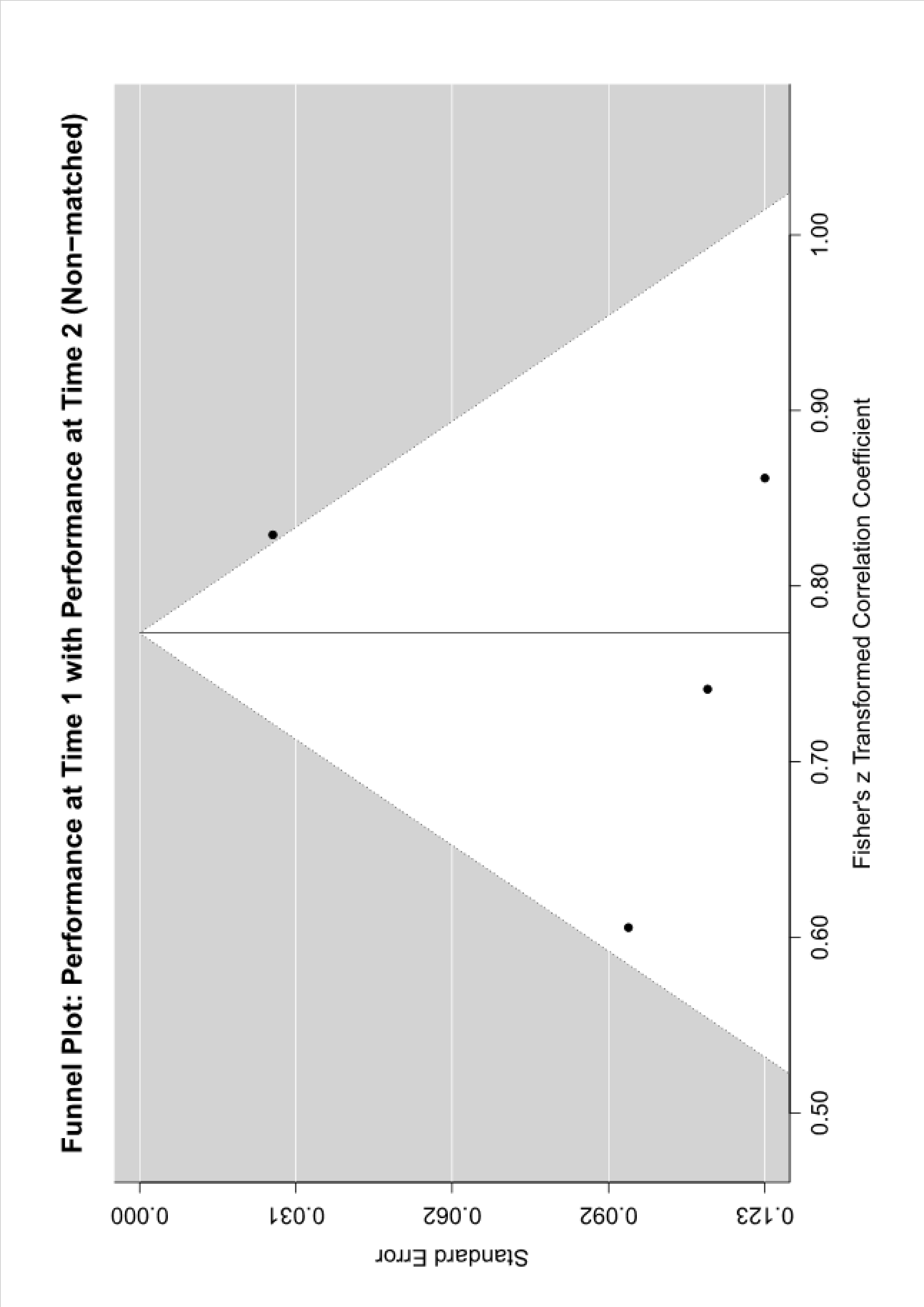

Forest Plot: Self−efficacy at Time 2 with Performance at Time 2 (Matched)
n
Meier et al., 1984
71
0.09 [−0.15 , 0.33]
Galyon et al., 2012
213
0.22 [0.08 , 0.35]
Mone et al., 1995
215
0.24 [0.11 , 0.38]
Bonne et al., 2016
50
0.32 [0.03 , 0.61]
Vancouver et al., 2006
65
0.33 [0.08 , 0.58]
Mone, 1994
252
0.33 [0.21 , 0.46]
Finney et al., 2003
103
0.47 [0.28 , 0.67]
RE Model
0.28 [0.21 , 0.36]
−0.20
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient


Funnel Plot: Self−efficacy at Time 1 with Performance at Time 1 (Non−matched)
Standard Error
0.000
0.031
0.062
0.092
0.123
0.10
0.20
0.30
0.40
0.50
0.60
Fisher's z Transformed Correlation Coefficient

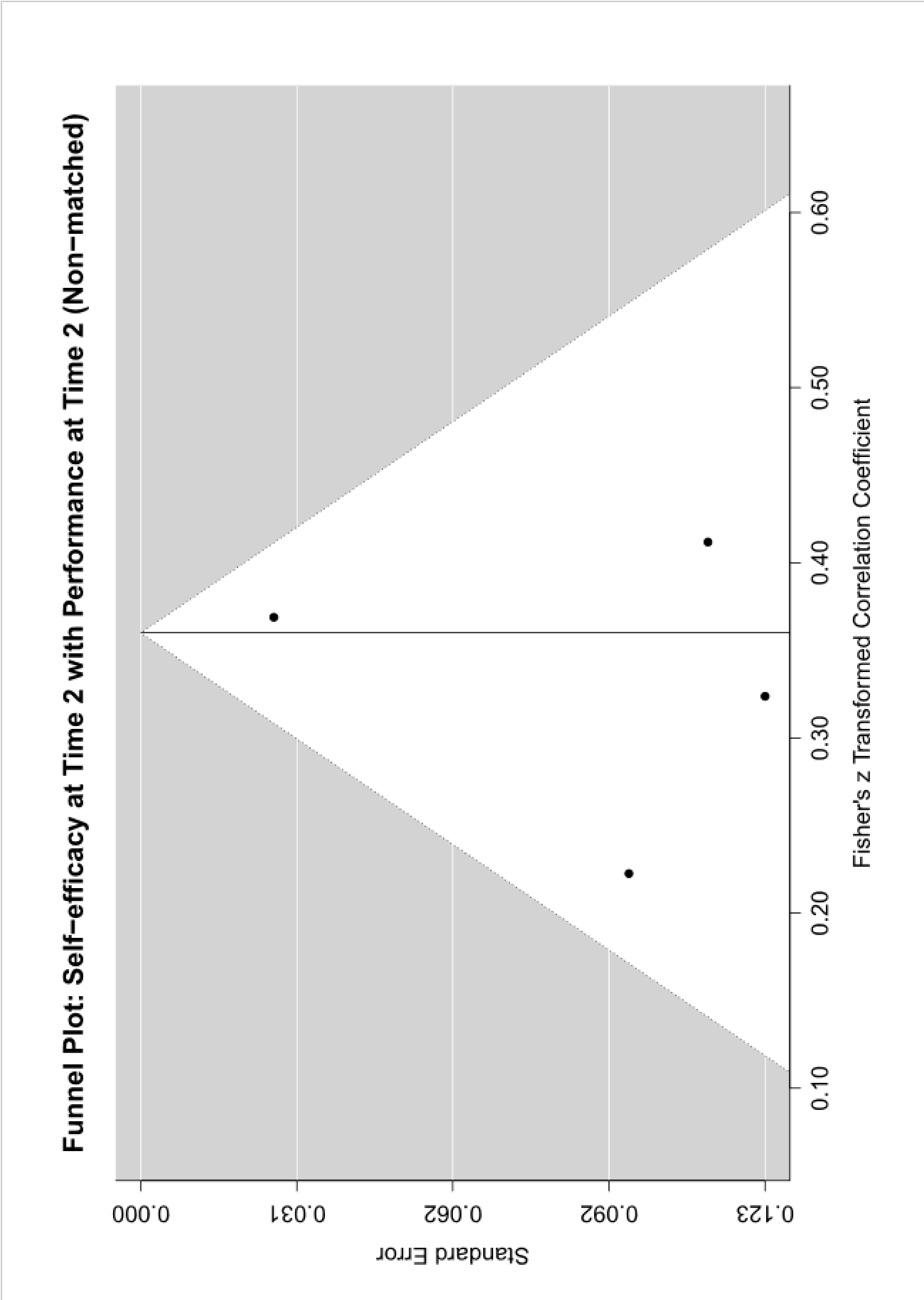

Funnel Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2 (Non−matched)
Standard Error
0.000
0.031
0.062
0.092
0.123
Fisher's z Transformed Correlation Coefficient
0.40
0.50
0.60
0.70
0.80
0.90
1.00
1.10

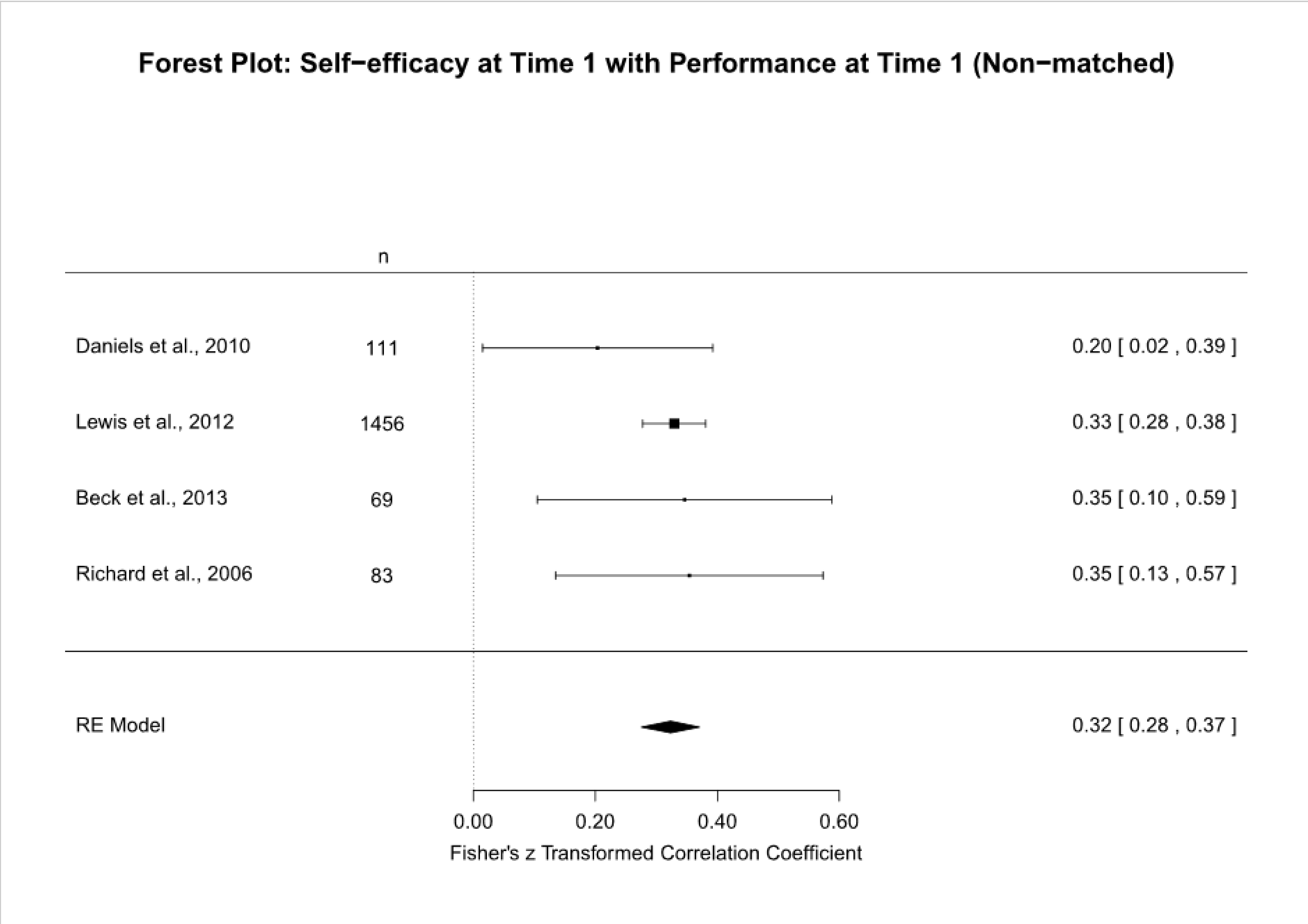

Funnel Plot: Self−efficacy at Time 1 with Performance at Time 2 (Non−matched)
Standard Error
0.000
0.031
0.062
0.092
0.123
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50

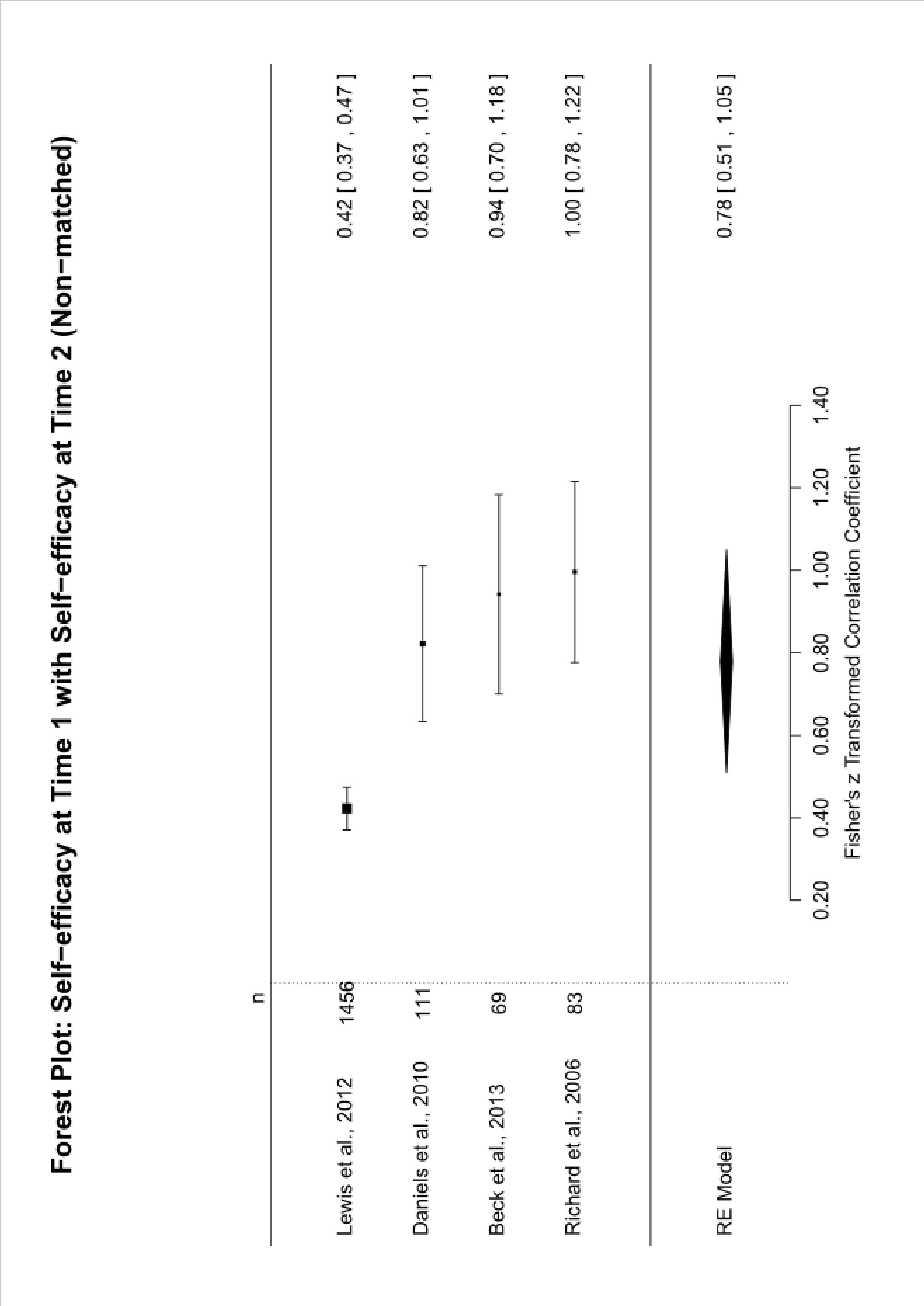

Funnel Plot: Self−efficacy at Time 2 with Performance at Time 1 (Non−matched)
Standard Error
0.000
0.031
0.062
0.092
0.123
Fisher's z Transformed Correlation Coefficient
0.20
0.30
0.40
0.50
0.60
0.70

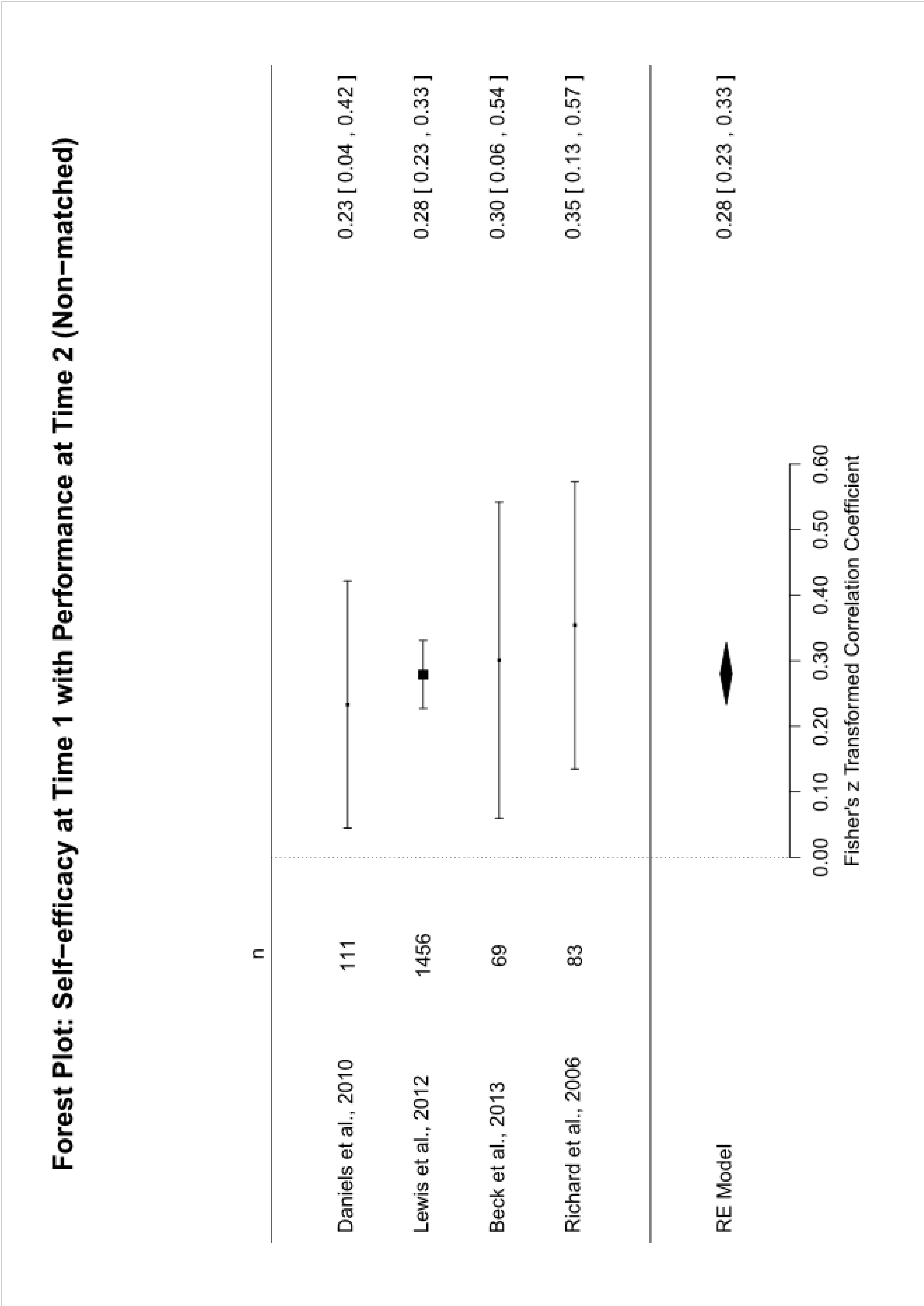


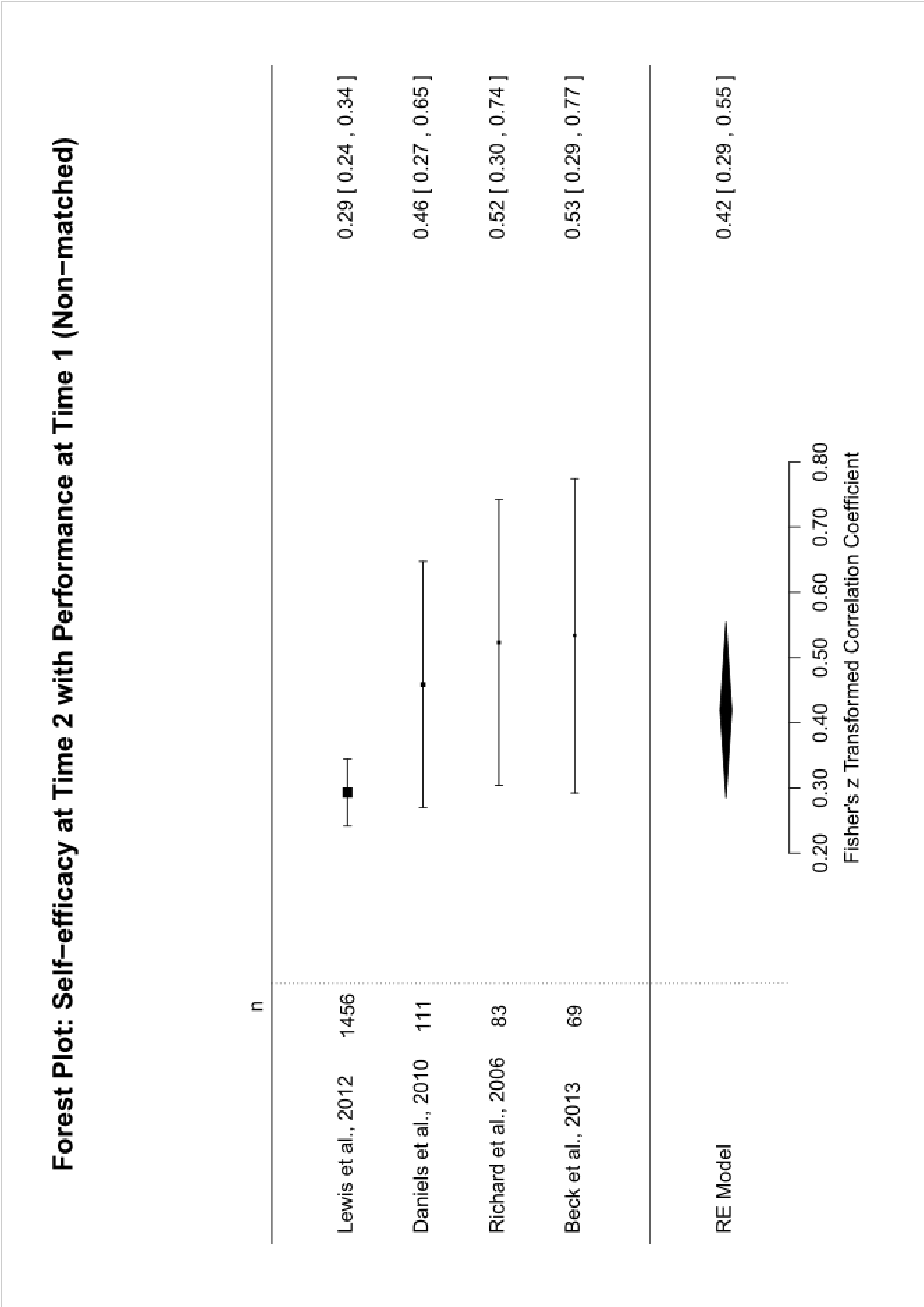
Funnel Plot: Performance at Time 1 with Performance at Time 2 (Non−matched)
Standard Error
0.000
0.031
0.062
0.092
0.123
Fisher's z Transformed Correlation Coefficient
0.50
0.60
0.70
0.80
0.90
1.00

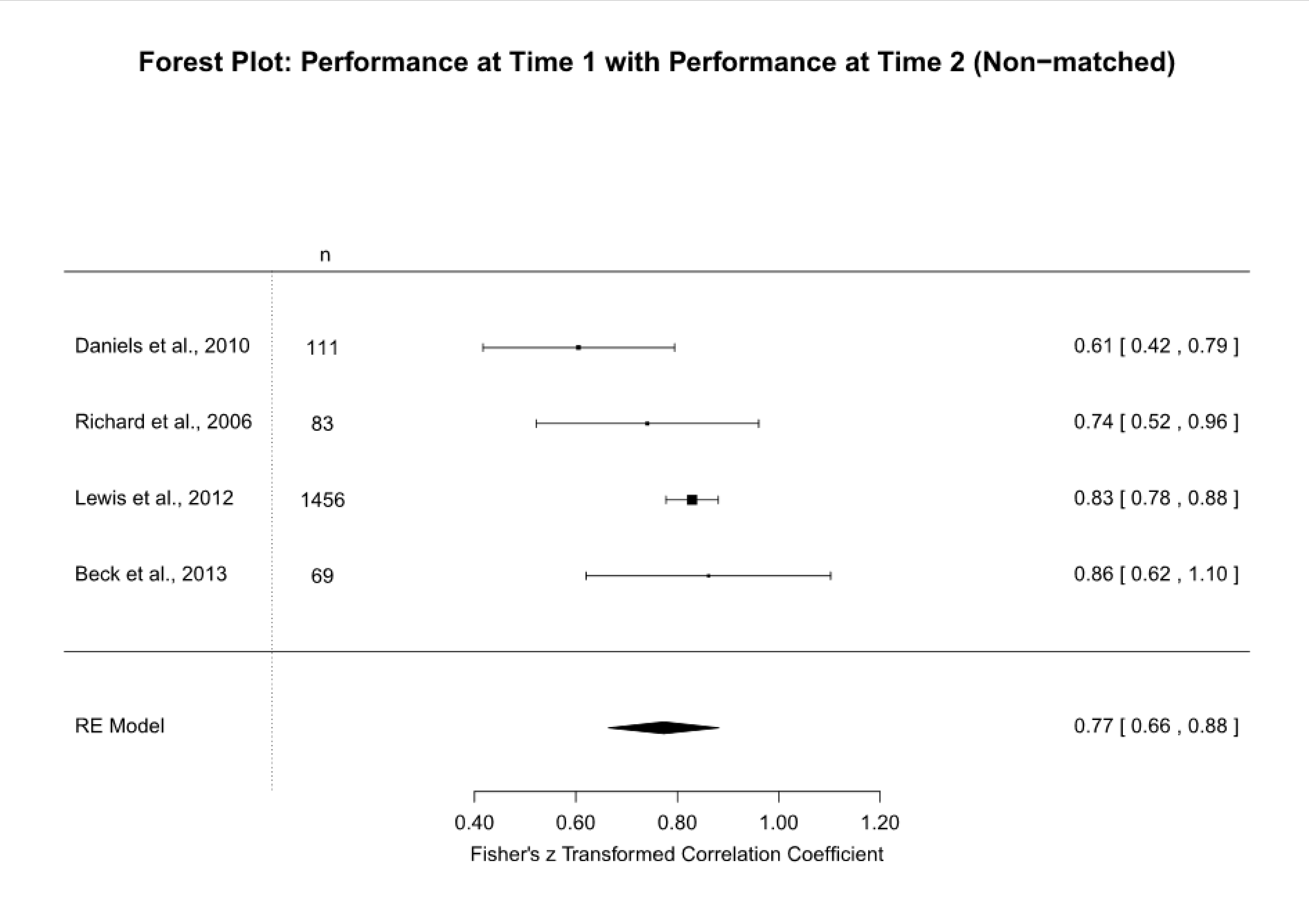

Funnel Plot: Self−efficacy at Time 2 with Performance at Time 2 (Non−matched)
Standard Error
0.000
0.031
0.062
0.092
0.123
Fisher's z Transformed Correlation Coefficient
0.10
0.20
0.30
0.40
0.50
0.60

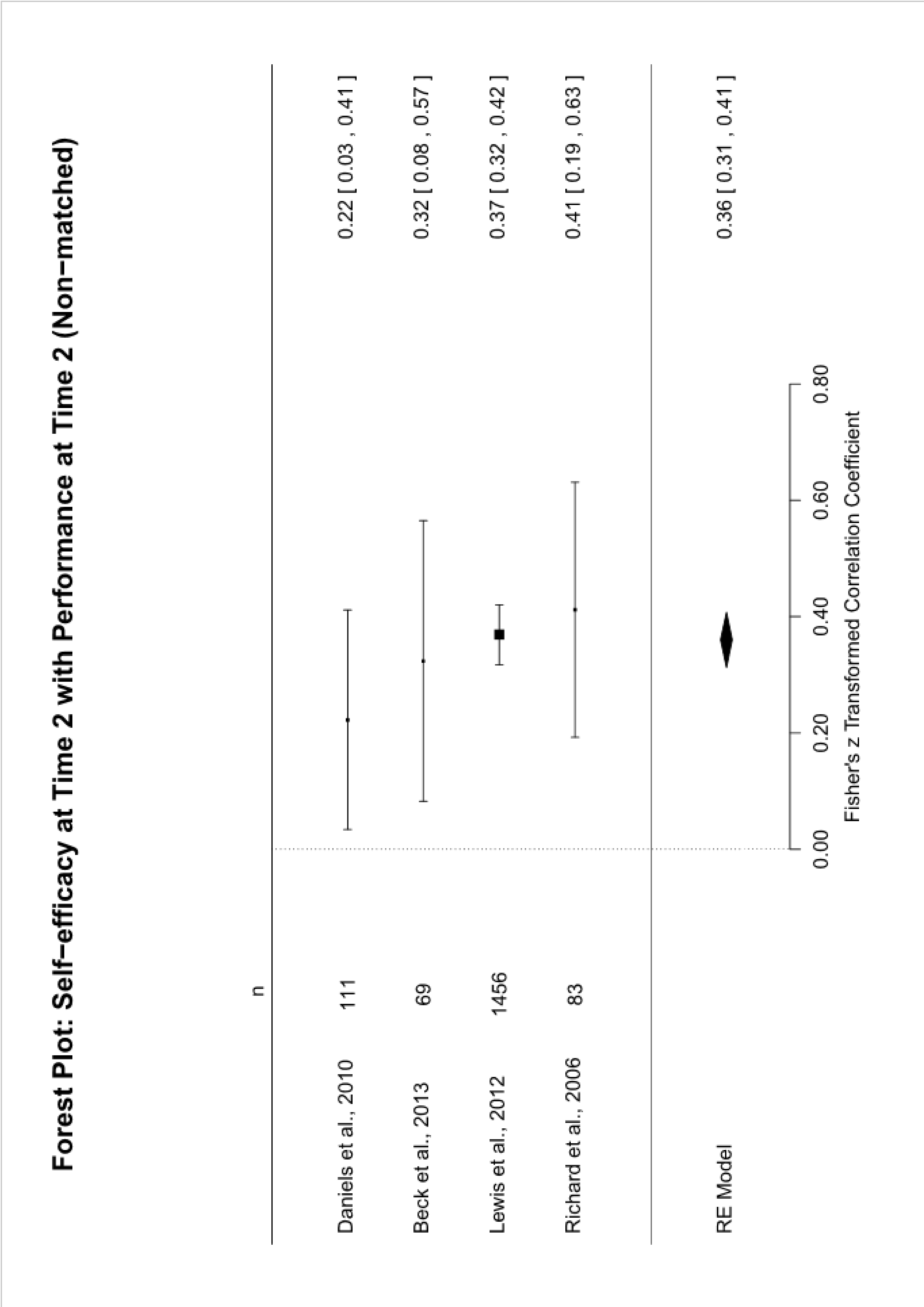

Forest Plot: Self−efficacy at Time 1 with Performance at Time 1 (Non−matched)
n
Daniels et al., 2010
111
0.20 [0.02 , 0.39]
Lewis et al., 2012
1456
0.33 [0.28 , 0.38]
Beck et al., 2013
69
0.35 [0.10 , 0.59]
Richard et al., 2006
83
0.35 [0.13 , 0.57]
RE Model
0.32 [0.28 , 0.37]
0.00
0.20
0.40
0.60
Fisher's z Transformed Correlation Coefficient

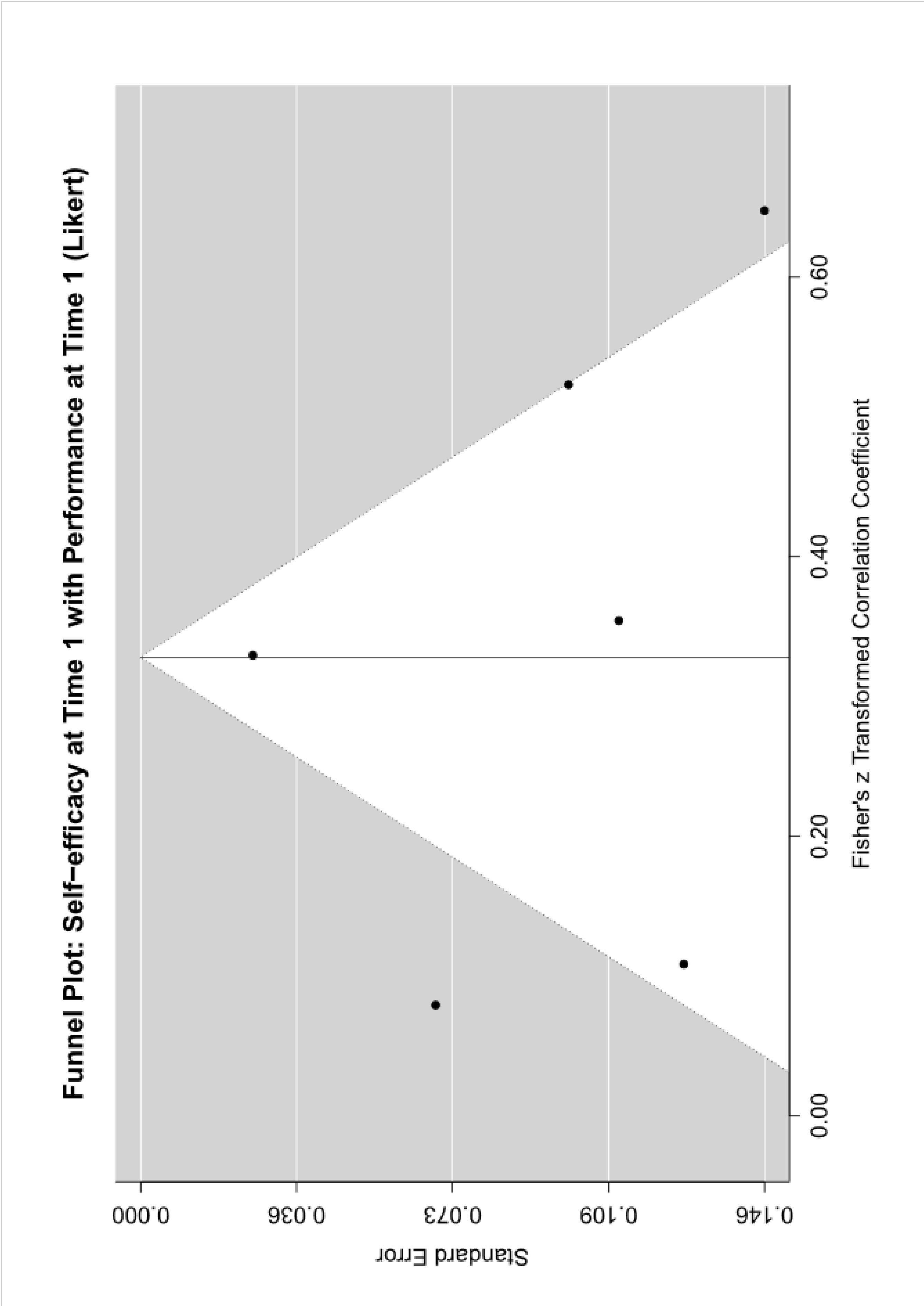

Forest Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2 (Non−matched)
n
Lewis et al., 2012
1456
0.42 [0.37 , 0.47]
Daniels et al., 2010
111
0.82 [0.63 , 1.01]
Beck et al., 2013
69
0.94 [0.70 , 1.18]
Richard et al., 2006
83
1.00 [0.78 , 1.22]
RE Model
0.78 [0.51 , 1.05]
0.20
0.40
0.60
0.80
1.00
1.20
1.40
Fisher's z Transformed Correlation Coefficient

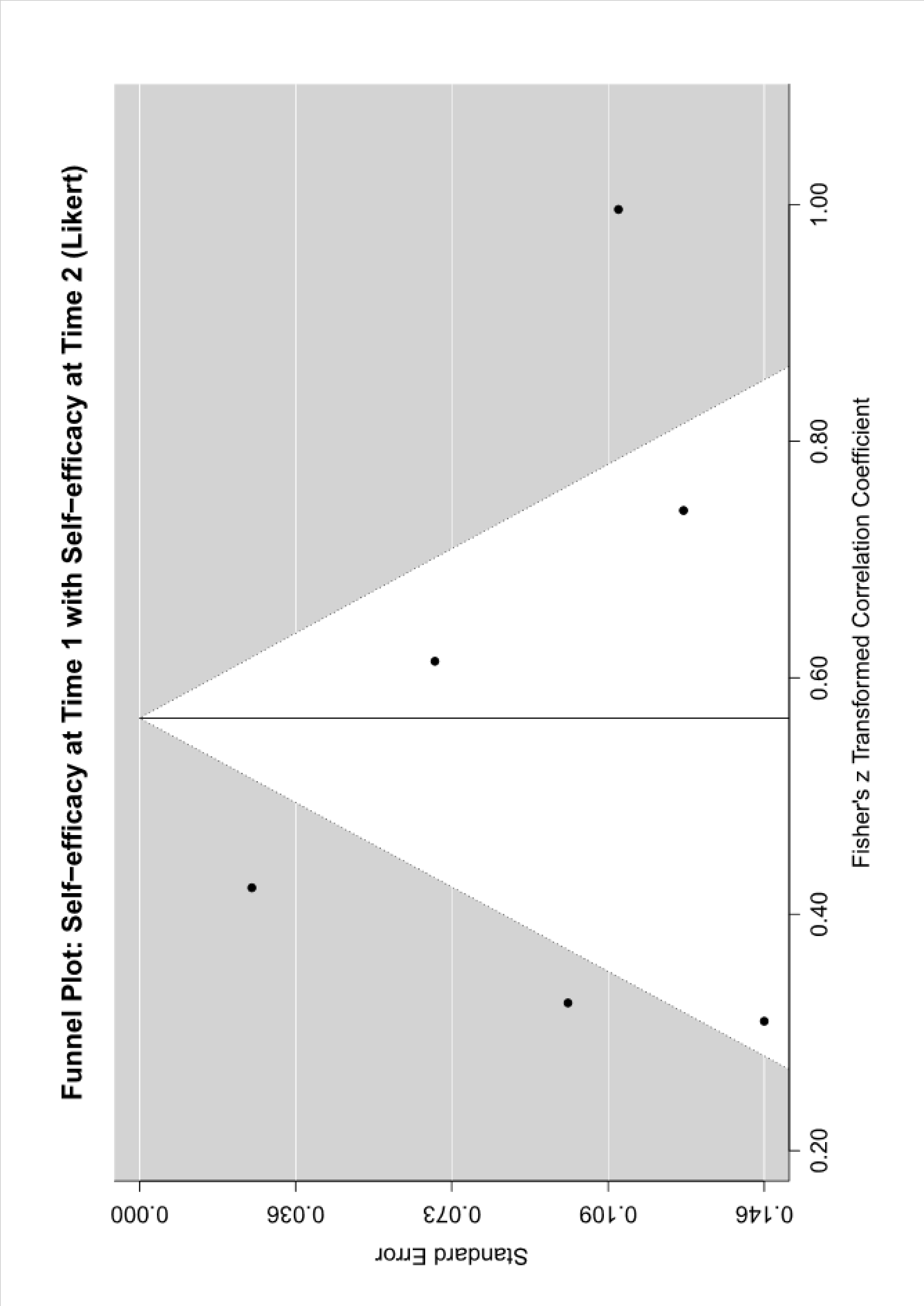

Forest Plot: Self−efficacy at Time 1 with Performance at Time 2 (Non−matched)
n
Daniels et al., 2010
111
0.23 [0.04 , 0.42]
Lewis et al., 2012
1456
0.28 [0.23 , 0.33]
Beck et al., 2013
69
0.30 [0.06 , 0.54]
Richard et al., 2006
83
0.35 [0.13 , 0.57]
RE Model
0.28 [0.23 , 0.33]
0.00
0.10
0.20
0.30
0.40
0.50
0.60
Fisher's z Transformed Correlation Coefficient

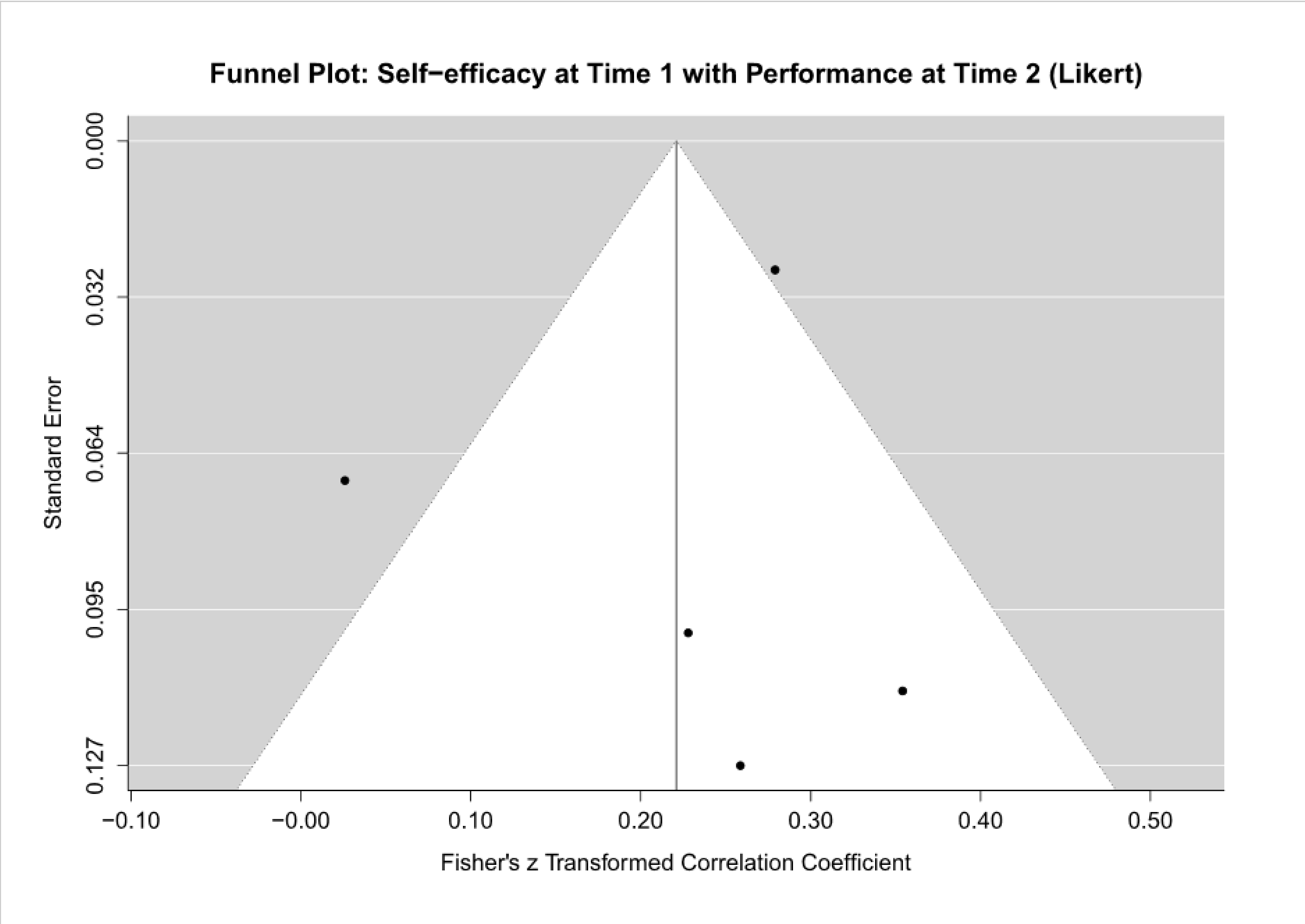

Forest Plot: Self−efficacy at Time 2 with Performance at Time 1 (Non−matched)
n
Lewis et al., 2012 1456 0.29 [0.24 , 0.34]
Daniels et al., 2010 111 0.46 [0.27 , 0.65]
Richard et al., 2006 83 0.52 [0.30 , 0.74]
Beck et al., 2013 69 0.53 [0.29 , 0.77]
RE Model 0.42 [0.29 , 0.55]
0.20 0.30 0.40 0.50 0.60 0.70 0.80
Fisher's z Transformed Correlation Coefficient

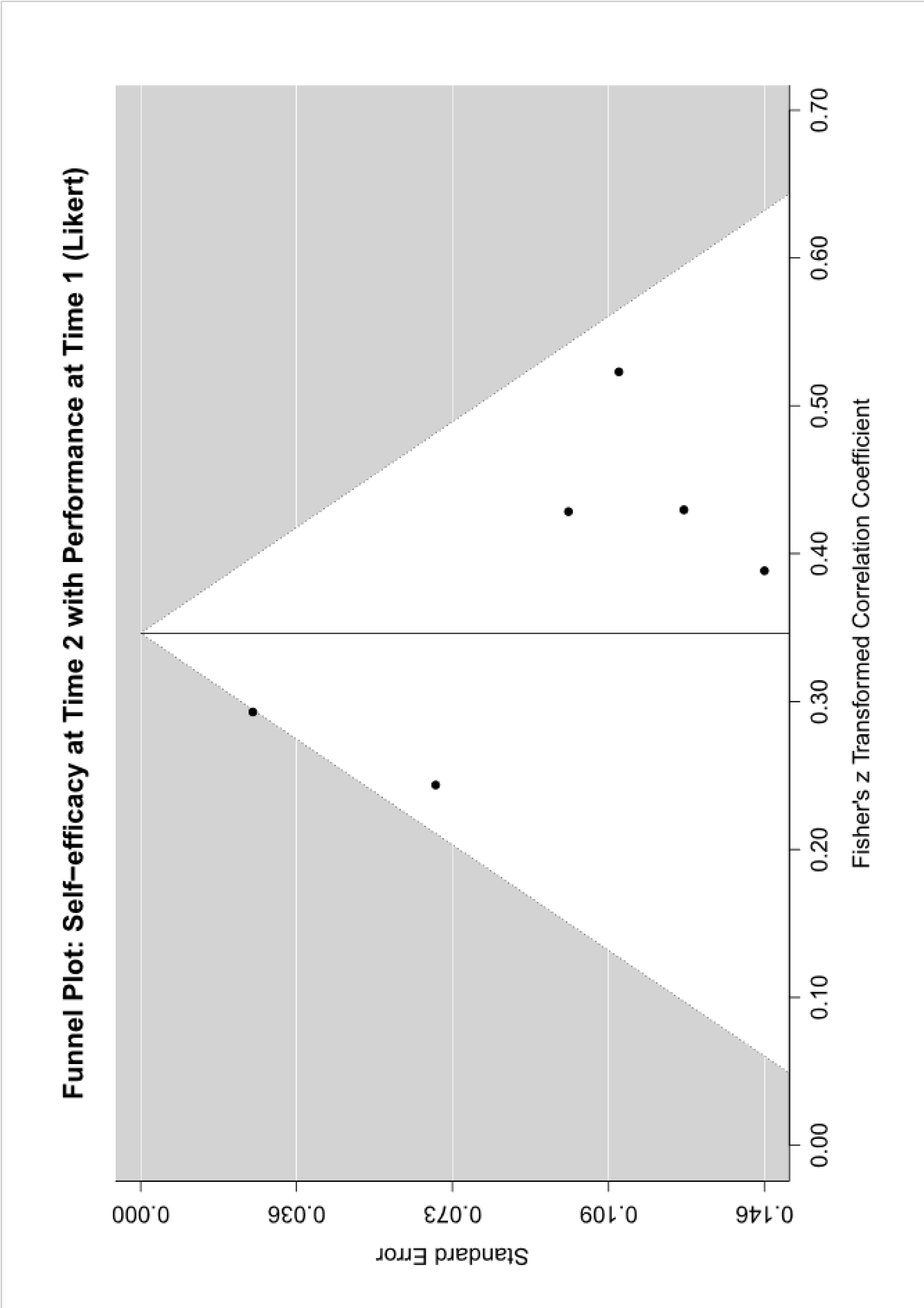

Forest Plot: Performance at Time 1 with Performance at Time 2 (Non−matched)
n
Daniels et al., 2010
111
0.61 [0.42 , 0.79]
Richard et al., 2006
83
0.74 [0.52 , 0.96]
Lewis et al., 2012
1456
0.83 [0.78 , 0.88]
Beck et al., 2013
69
0.86 [0.62 , 1.10]
RE Model
0.77 [0.66 , 0.88]
0.40
0.60
0.80
1.00
1.20
Fisher's z Transformed Correlation Coefficient

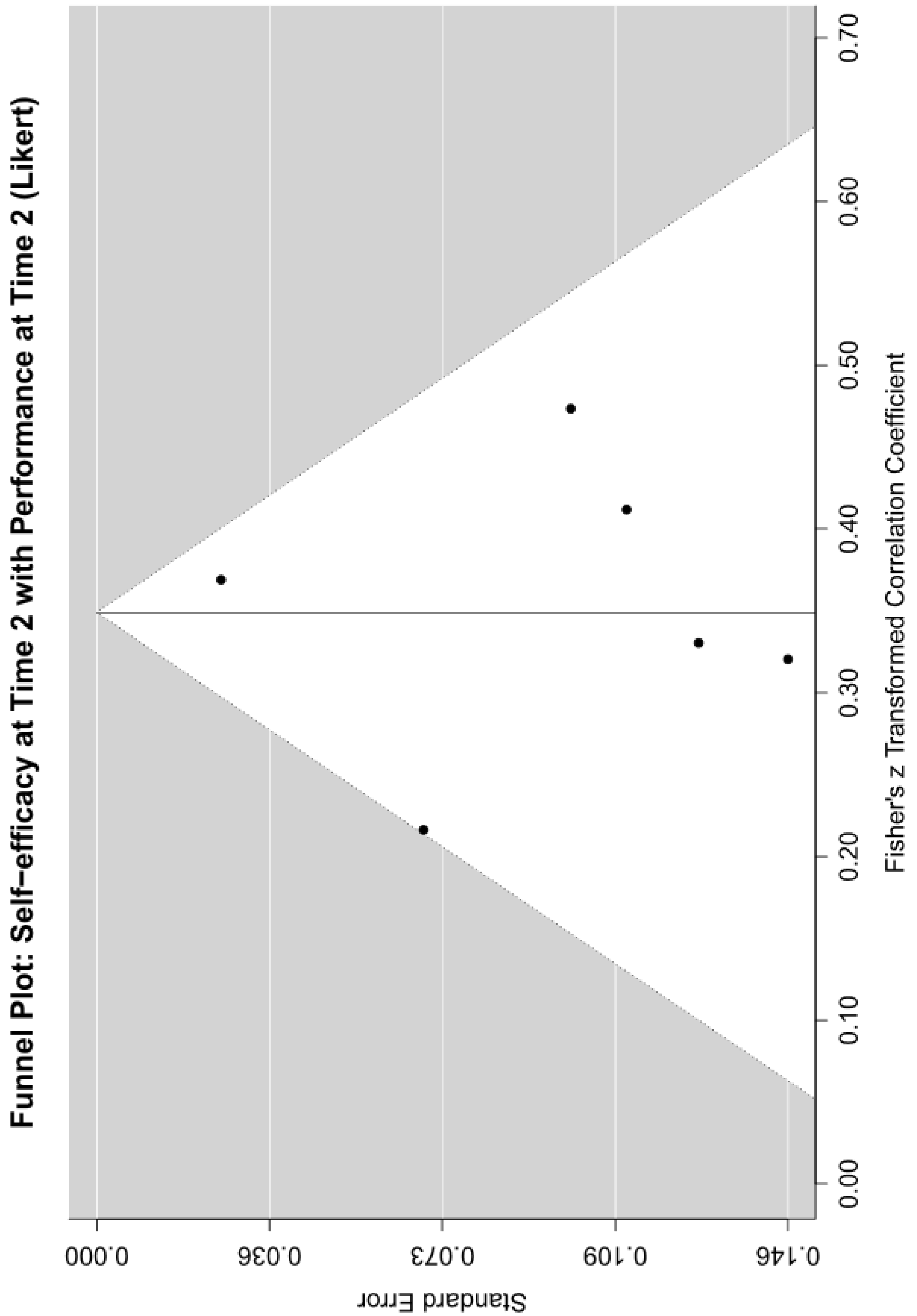

Forest Plot: Self−efficacy at Time 2 with Performance at Time 2 (Non−matched)
n
Daniels et al., 2010
111
0.22 [0.03 , 0.41]
Beck et al., 2013
69
0.32 [0.08 , 0.57]
Lewis et al., 2012
1456
0.37 [0.32 , 0.42]
Richard et al., 2006
83
0.41 [0.19 , 0.63]
RE Model
0.36 [0.31 , 0.41]
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient

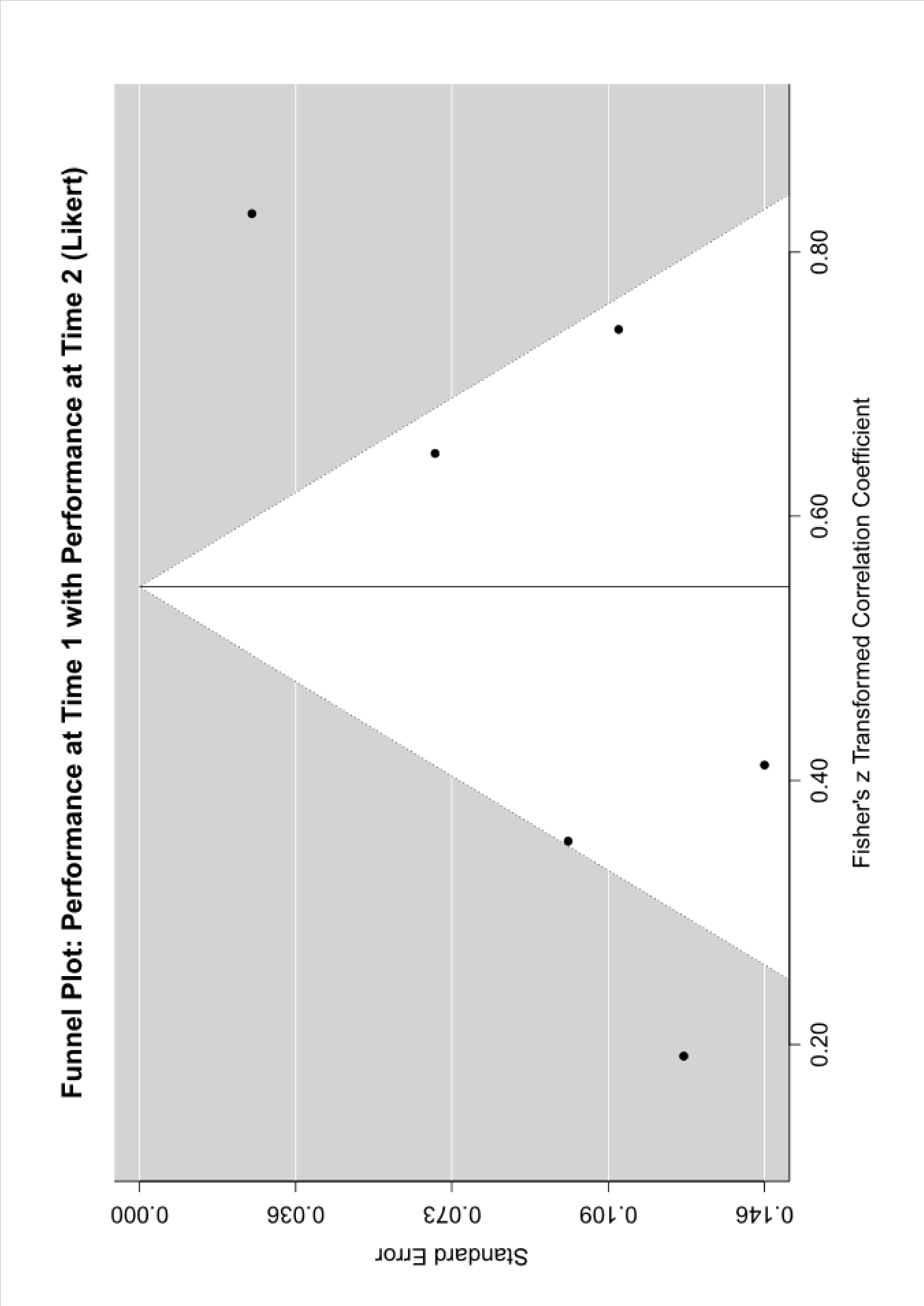

Funnel Plot: Self−efficacy at Time 1 with Performance at Time 1 (Likert)
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.00
0.20
0.40
0.60

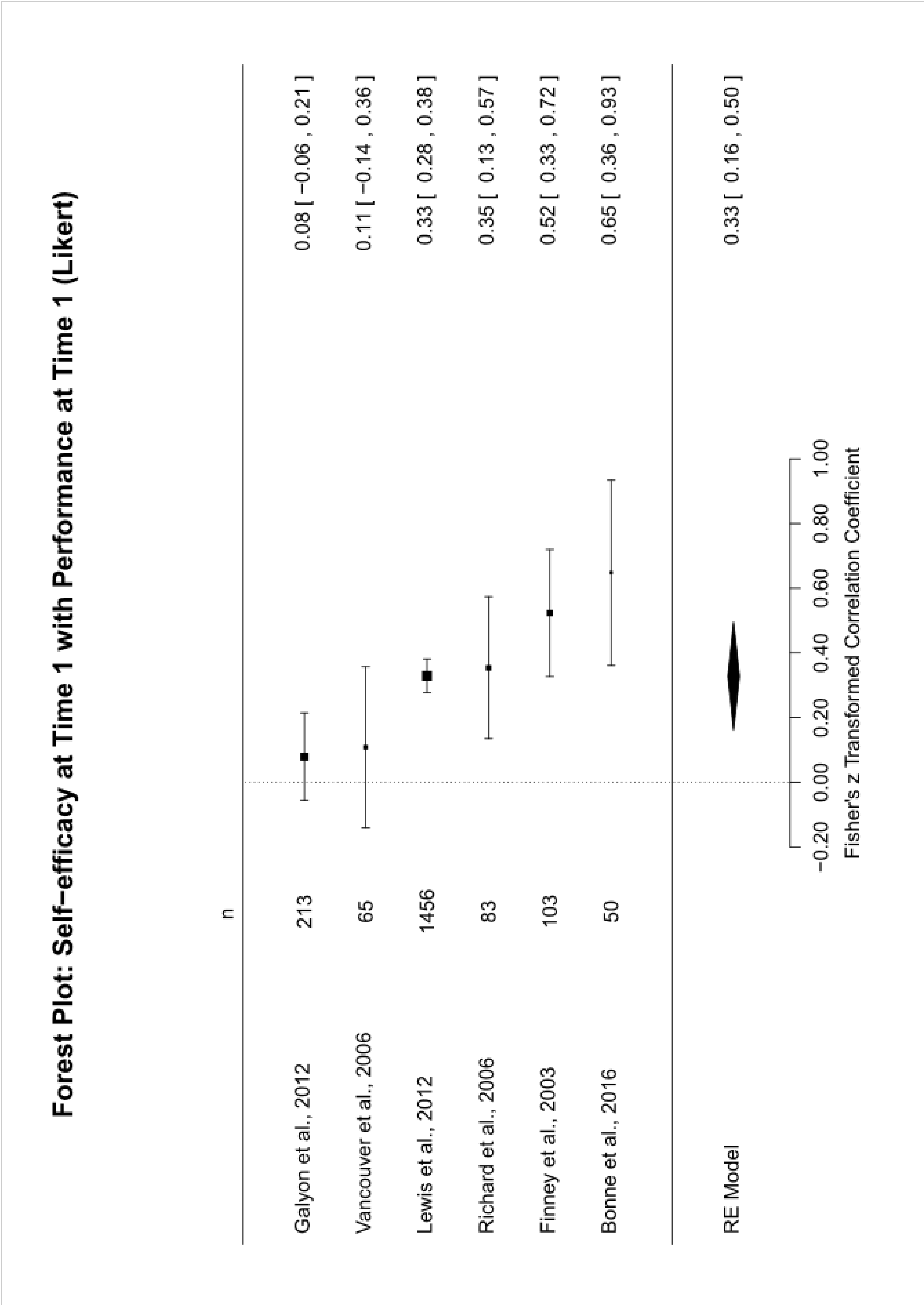

Funnel Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2 (Likert)
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.20
0.40
0.60
0.80
1.00

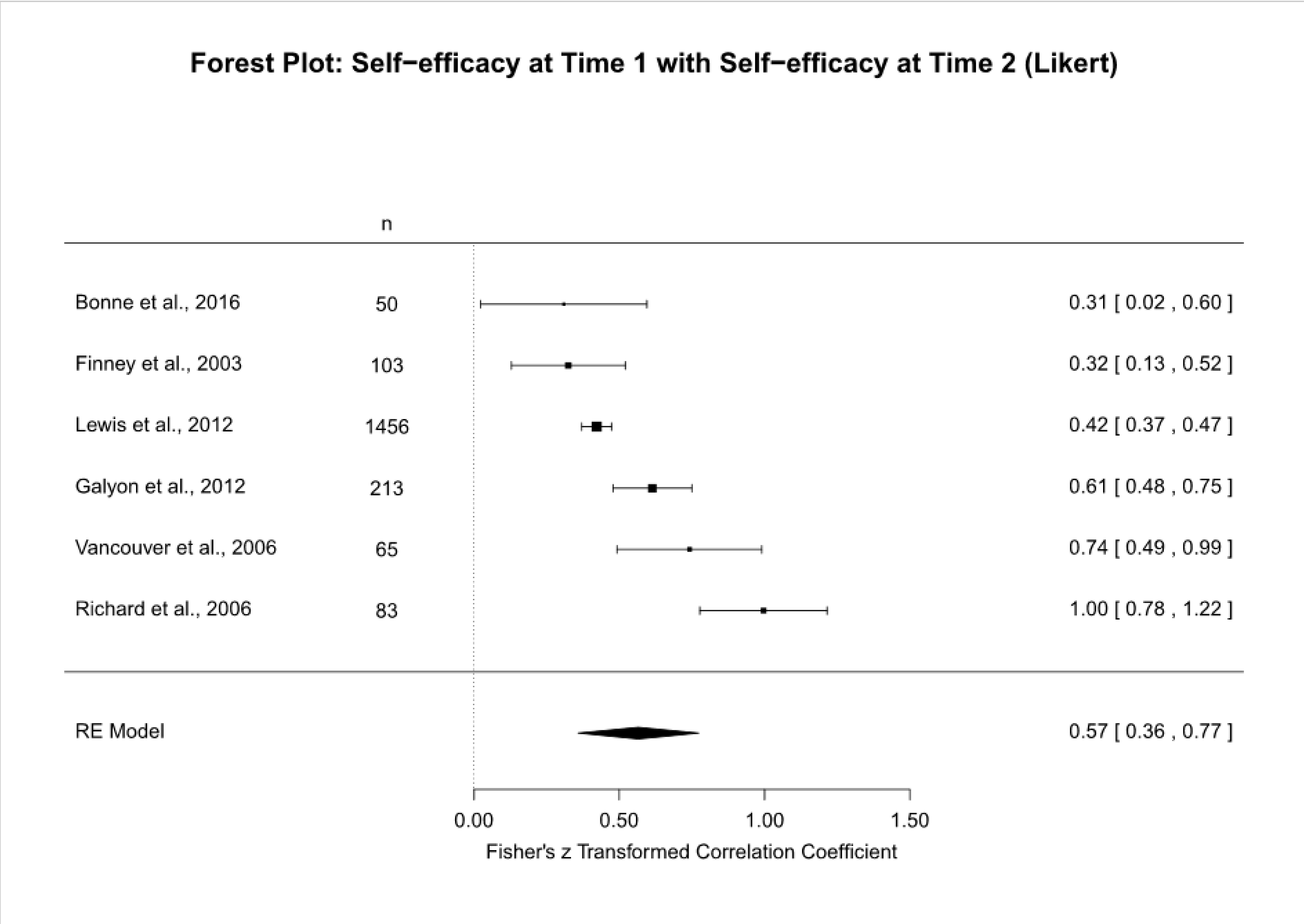

Funnel Plot: Self−efficacy at Time 1 with Performance at Time 2 (Likert)
Standard Error
0.000
0.032
0.064
0.095
0.127
−0.10
−0.00
0.10
0.20
0.30
0.40
0.50
Fisher's z Transformed Correlation Coefficient

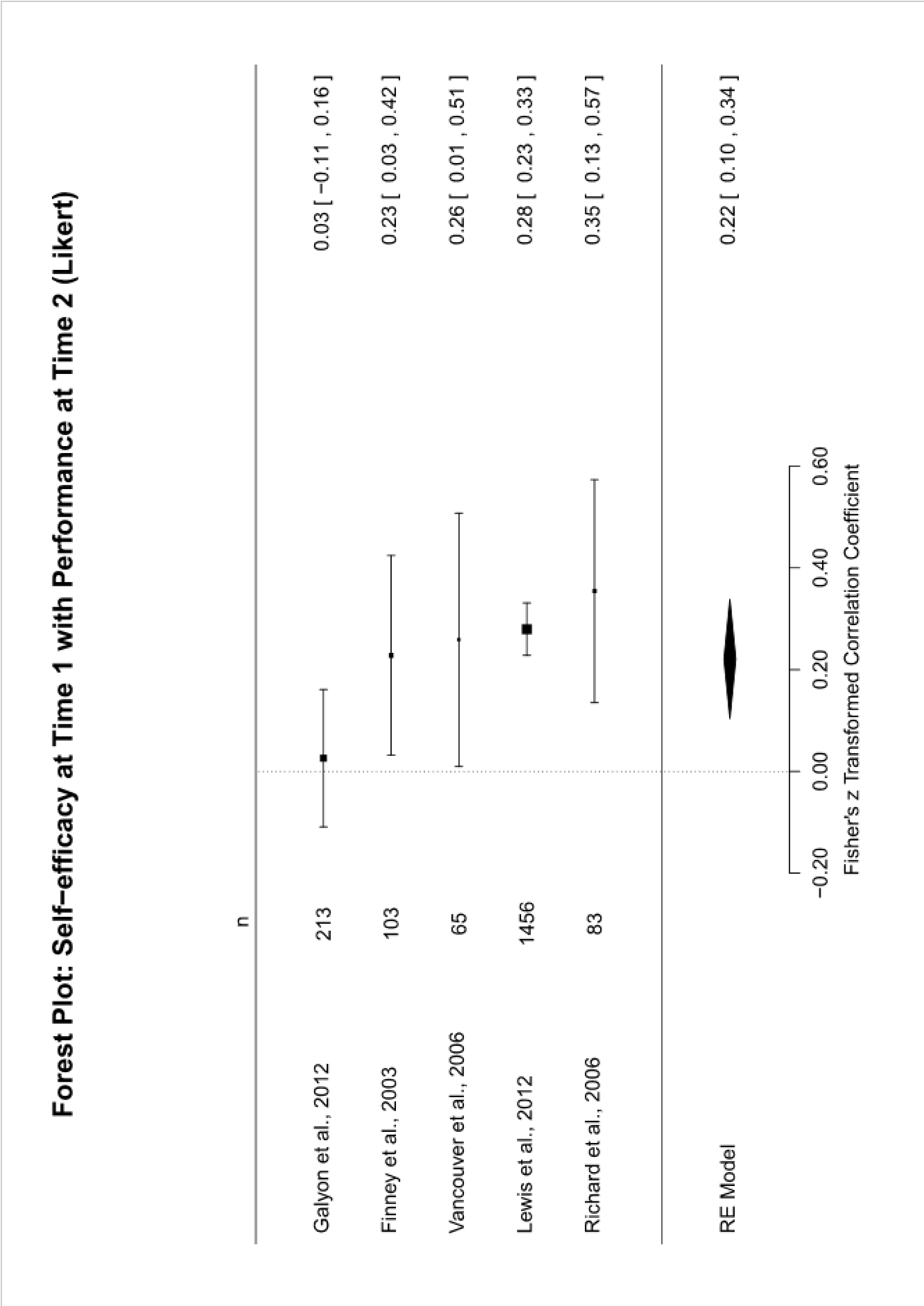

Funnel Plot: Self−efficacy at Time 2 with Performance at Time 1 (Likert)
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70

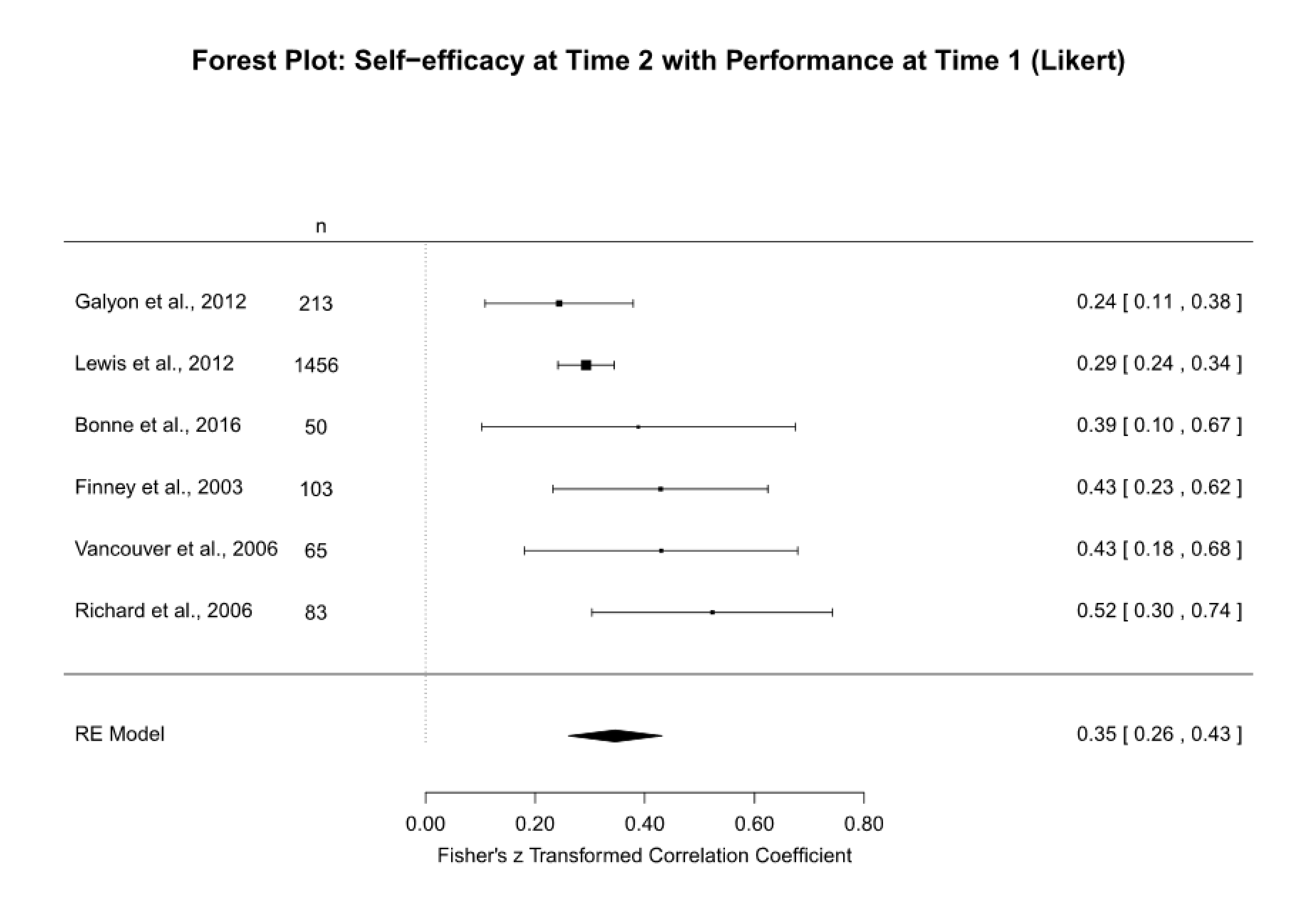

Funnel Plot: Self−efficacy at Time 2 with Performance at Time 2 (Likert)
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70

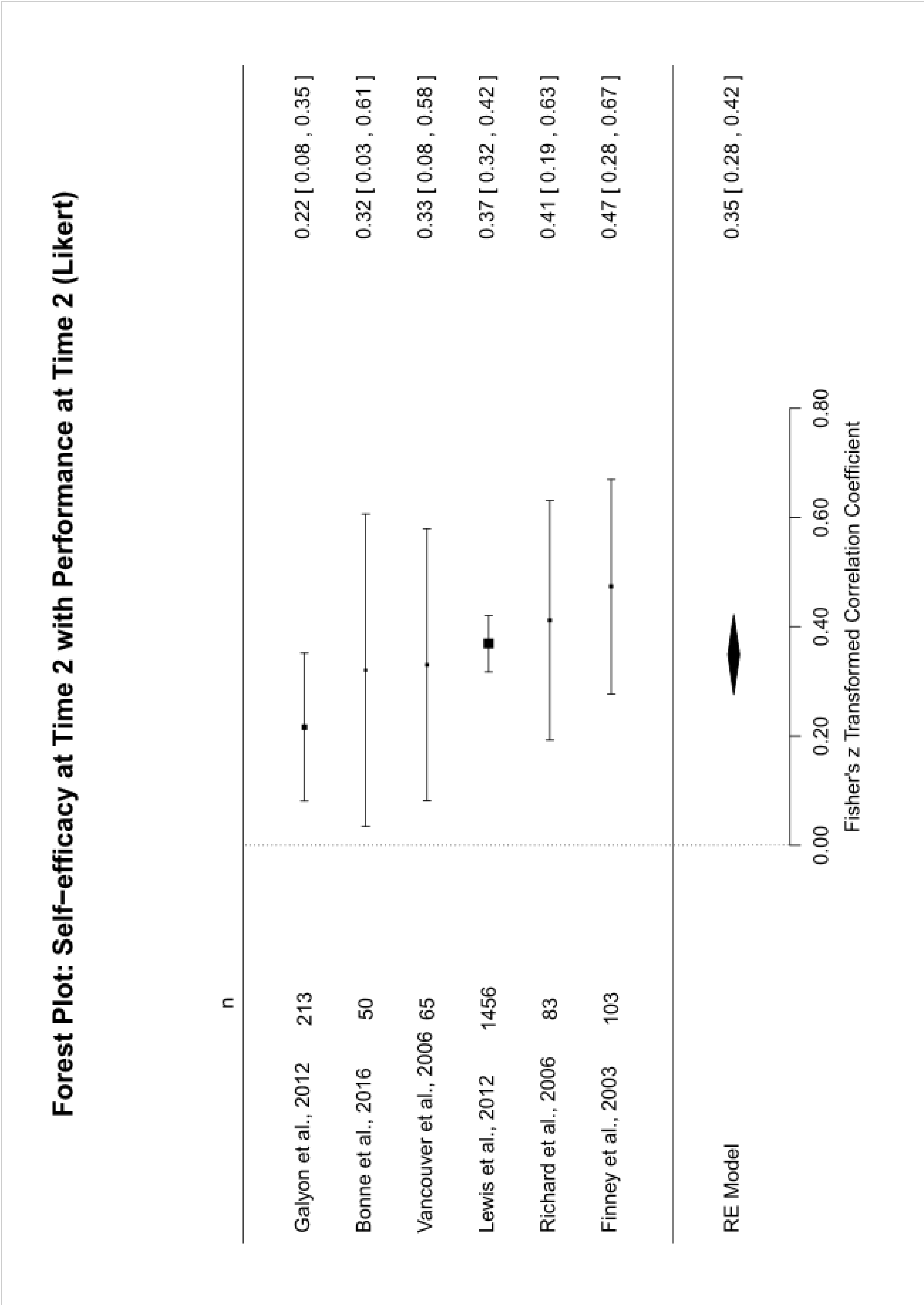

Funnel Plot: Performance at Time 1 with Performance at Time 2 (Likert)
Standard Error
0.000
0.036
0.073
0.109
0.146
Fisher's z Transformed Correlation Coefficient
0.20
0.40
0.60
0.80

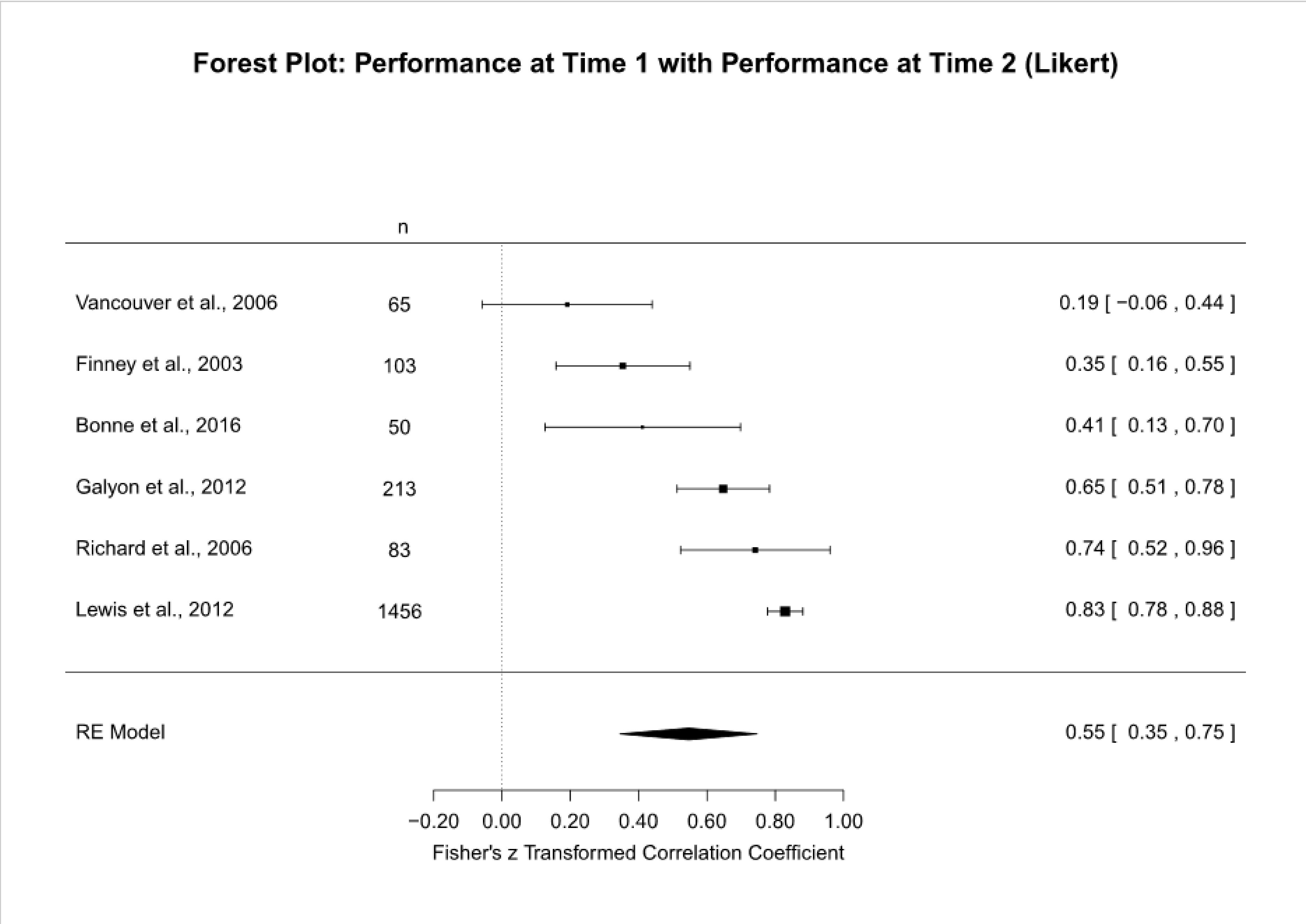

Forest Plot: Self−efficacy at Time 1 with Performance at Time 1 (Likert)
n
Galyon et al., 2012
213
0.08 [−0.06 , 0.21]
Vancouver et al., 2006
65
0.11 [−0.14 , 0.36]
Lewis et al., 2012
1456
0.33 [0.28 , 0.38]
Richard et al., 2006
83
0.35 [0.13 , 0.57]
Finney et al., 2003
103
0.52 [0.33 , 0.72]
Bonne et al., 2016
50
0.65 [0.36 , 0.93]
RE Model
0.33 [0.16 , 0.50]
−0.20 0.00 0.20 0.40 0.60 0.80 1.00
Fisher's z Transformed Correlation Coefficient


Forest Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2 (Likert)
n
Bonne et al., 2016
50
0.31 [0.02 , 0.60]
Finney et al., 2003
103
0.32 [0.13 , 0.52]
Lewis et al., 2012
1456
0.42 [0.37 , 0.47]
Galyon et al., 2012
213
0.61 [0.48 , 0.75]
Vancouver et al., 2006
65
0.74 [0.49 , 0.99]
Richard et al., 2006
83
1.00 [0.78 , 1.22]
RE Model
0.57 [0.36 , 0.77]
0.00
0.50
1.00
1.50
Fisher's z Transformed Correlation Coefficient

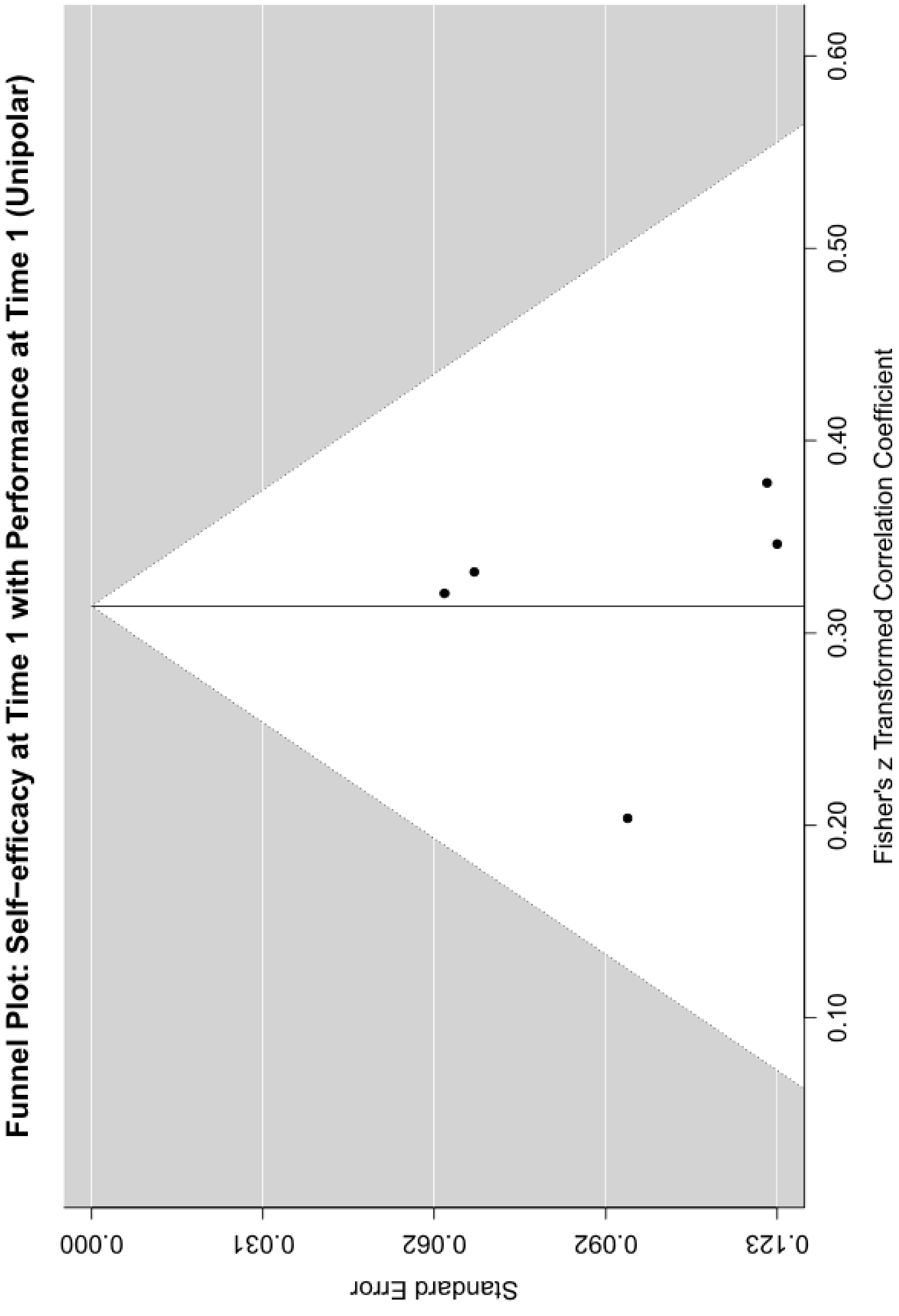

Forest Plot: Self–efficacy at Time 1 with Performance at Time 2 (Likert)
n
Galyon et al., 2012
213
0.03 [−0.11 , 0.16]
Finney et al., 2003
103
0.23 [0.03 , 0.42]
Vancouver et al., 2006
65
0.26 [0.01 , 0.51]
Lewis et al., 2012
1456
0.28 [0.23 , 0.33]
Richard et al., 2006
83
0.35 [0.13 , 0.57]
RE Model
0.22 [0.10 , 0.34]
−0.20
0.00
0.20
0.40
0.60
Fisher's z Transformed Correlation Coefficient


Forest Plot: Self−efficacy at Time 2 with Performance at Time 1 (Likert)
n
Galyon et al., 2012
213
0.24 [0.11 , 0.38]
Lewis et al., 2012
1456
0.29 [0.24 , 0.34]
Bonne et al., 2016
50
0.39 [0.10 , 0.67]
Finney et al., 2003
103
0.43 [0.23 , 0.62]
Vancouver et al., 2006
65
0.43 [0.18 , 0.68]
Richard et al., 2006
83
0.52 [0.30 , 0.74]
RE Model
0.35 [0.26 , 0.43]
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient

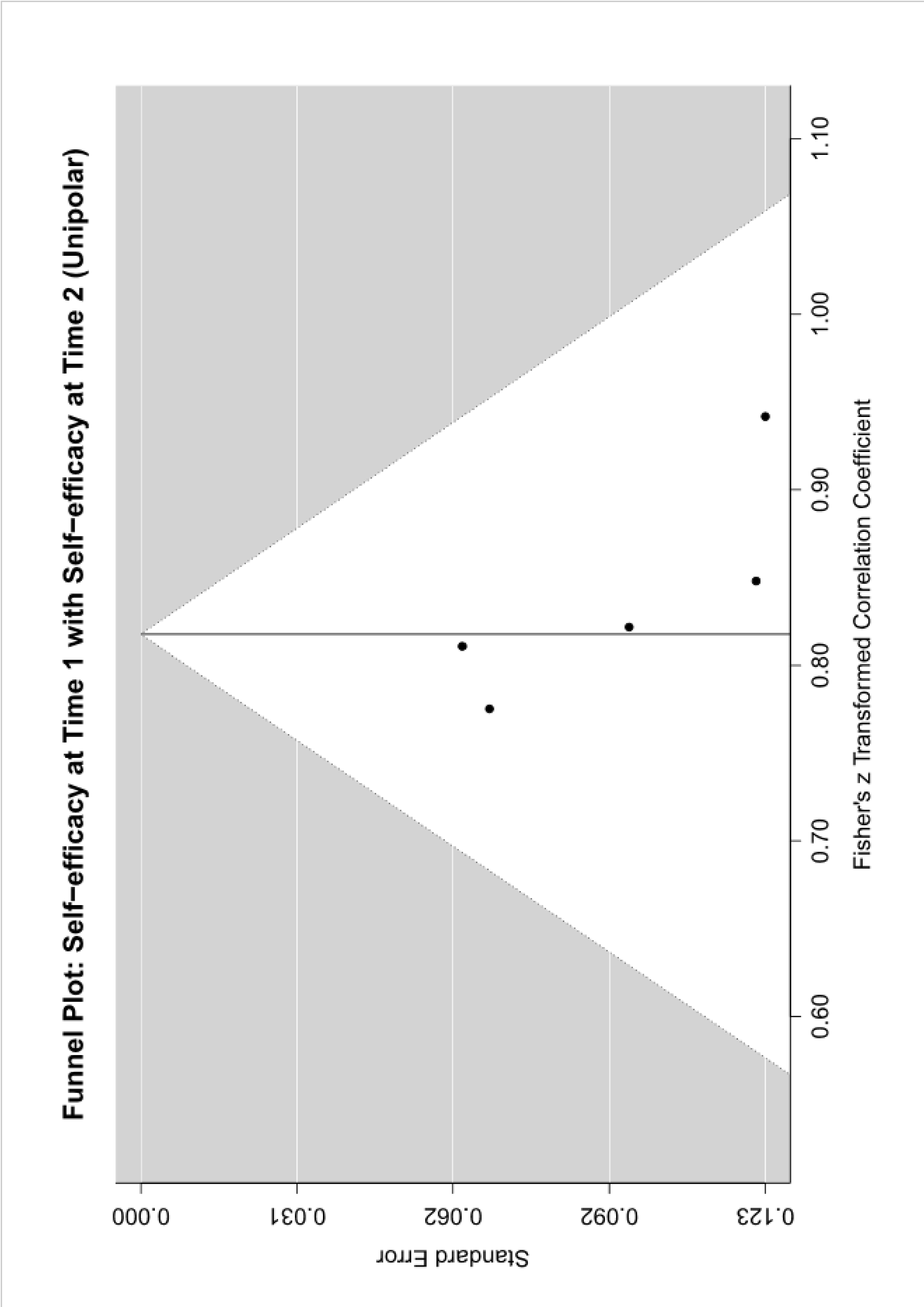

Forest Plot: Self–efficacy at Time 2 with Performance at Time 2 (Likert)
n
Galyon et al., 2012 213 0.22 [0.08 , 0.35]
Bonne et al., 2016 50 0.32 [0.03 , 0.61]
Vancouver et al., 2006 65 0.33 [0.08 , 0.58]
Lewis et al., 2012 1456 0.37 [0.32 , 0.42]
Richard et al., 2006 83 0.41 [0.19 , 0.63]
Finney et al., 2003 103 0.47 [0.28 , 0.67]
RE Model 0.35 [0.28 , 0.42]
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient

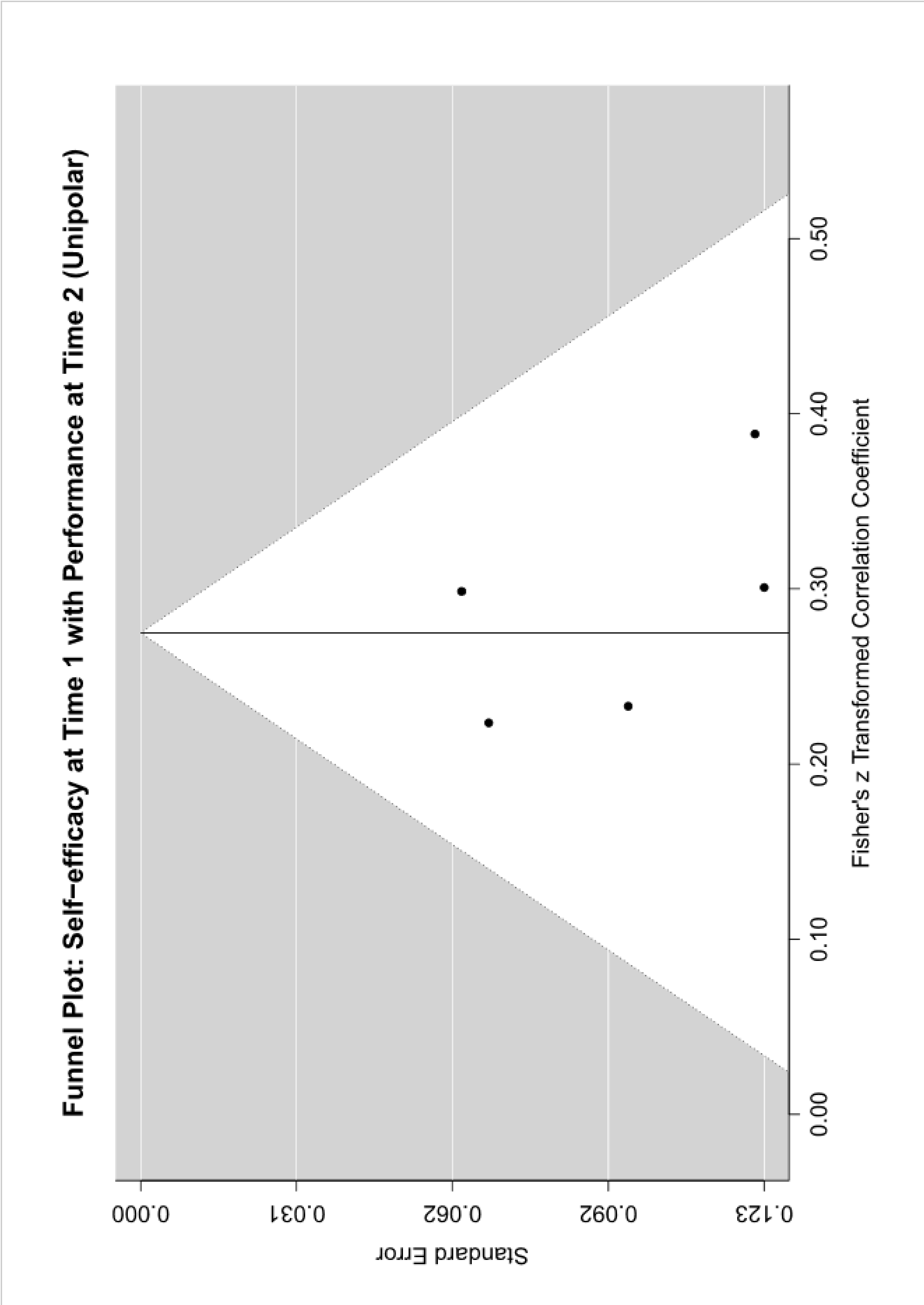


Forest Plot: Performance at Time 1 with Performance at Time 2 (Likert)
n
Vancouver et al., 2006
65
0.19 [−0.06 , 0.44]
Finney et al., 2003
103
0.35 [0.16 , 0.55]
Bonne et al., 2016
50
0.41 [0.13 , 0.70]
Galyon et al., 2012
213
0.65 [0.51 , 0.78]
Richard et al., 2006
83
0.74 [0.52 , 0.96]
Lewis et al., 2012
1456
0.83 [0.78 , 0.88]
RE Model
0.55 [0.35 , 0.75]
−0.20
0.00
0.20
0.40
0.60
0.80
1.00
Fisher's z Transformed Correlation Coefficient

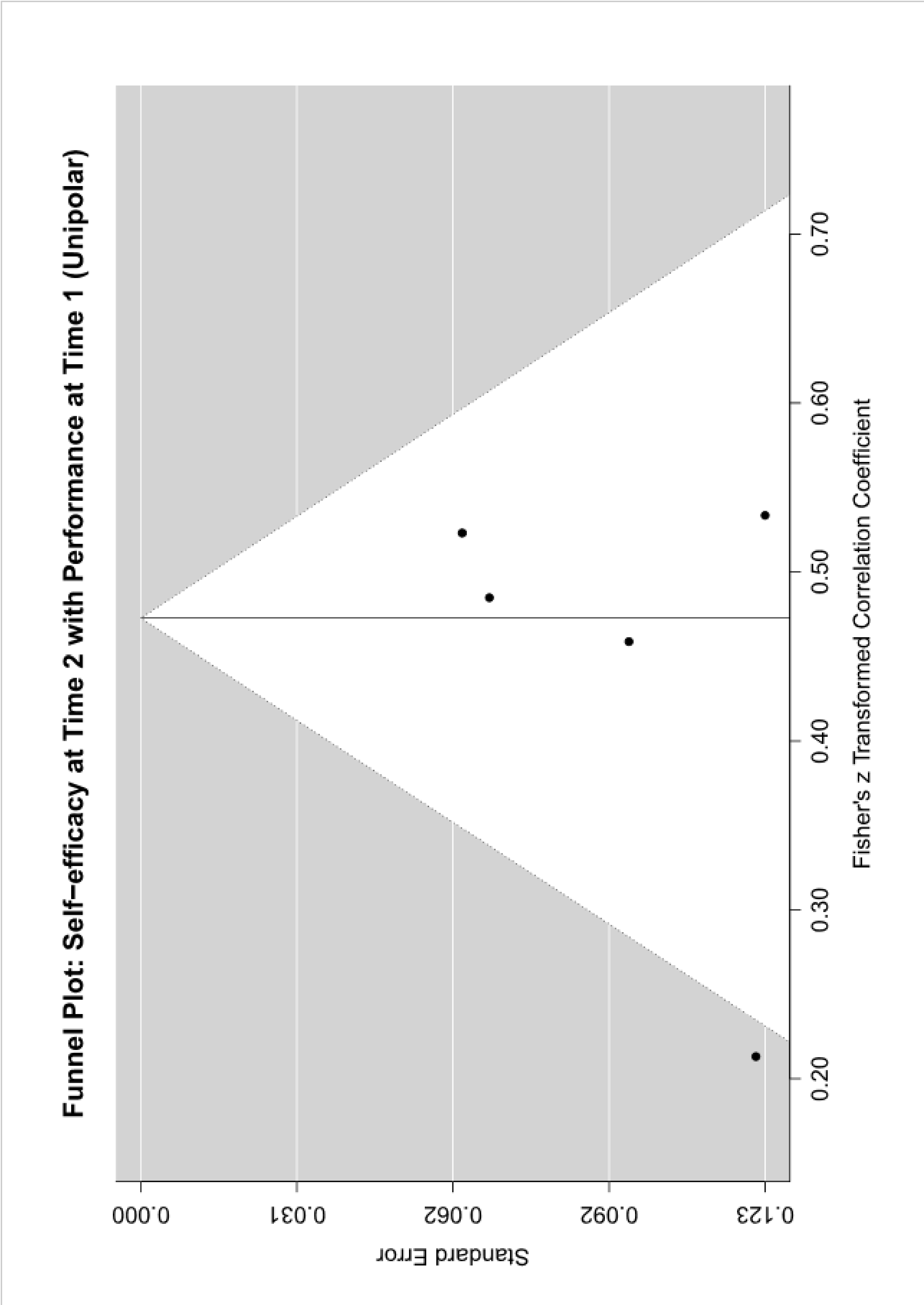


Funnel Plot: Self−efficacy at Time 1 with Performance at Time 1 (Unipolar)
Standard Error
0.000
0.031
0.062
0.092
0.123
Fisher's z Transformed Correlation Coefficient
0.10
0.20
0.30
0.40
0.50
0.60

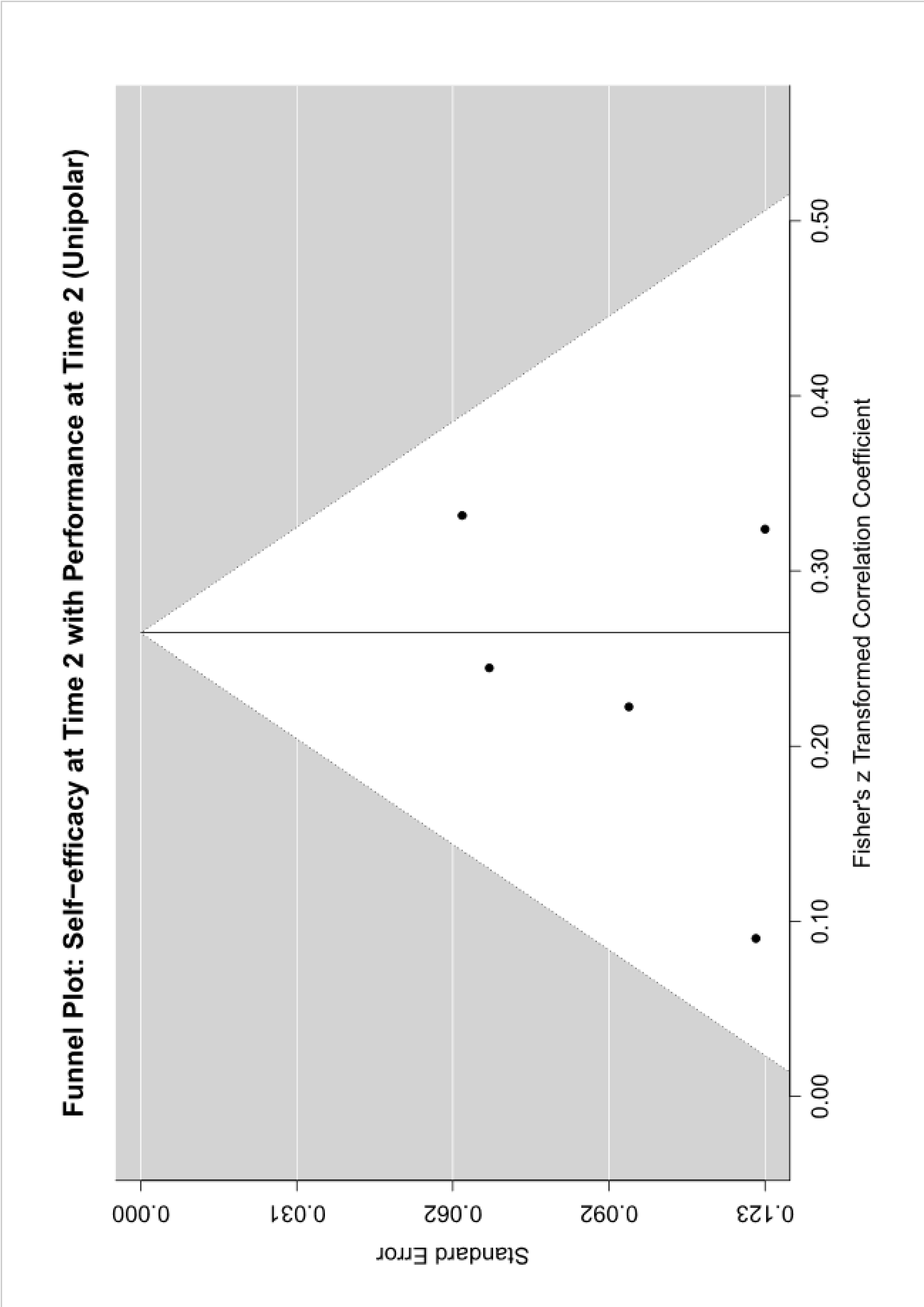

Funnel Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2 (Unipolar)
Standard Error
0.000
0.031
0.062
0.092
0.123
Fisher's z Transformed Correlation Coefficient
0.60
0.70
0.80
0.90
1.00
1.10

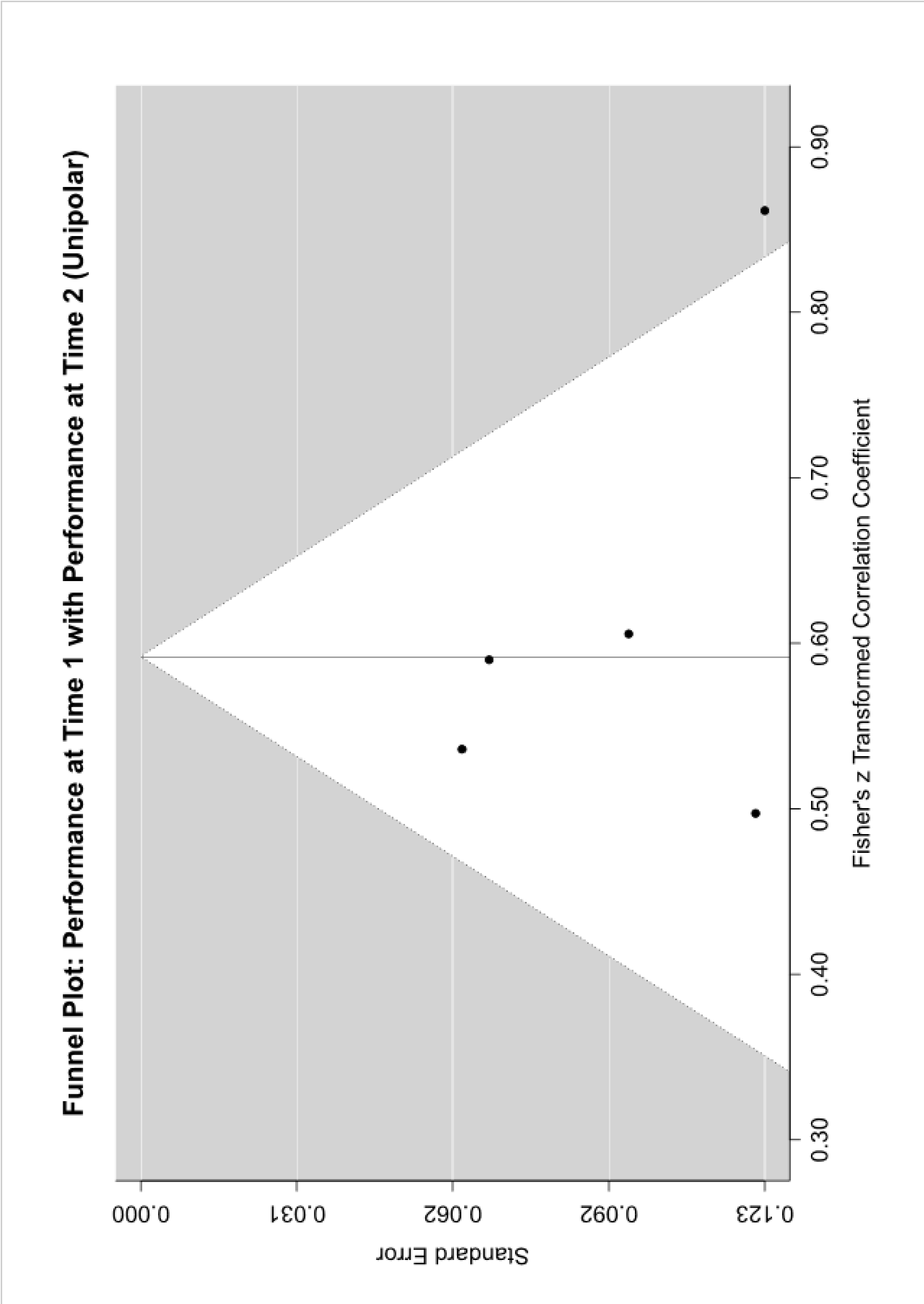


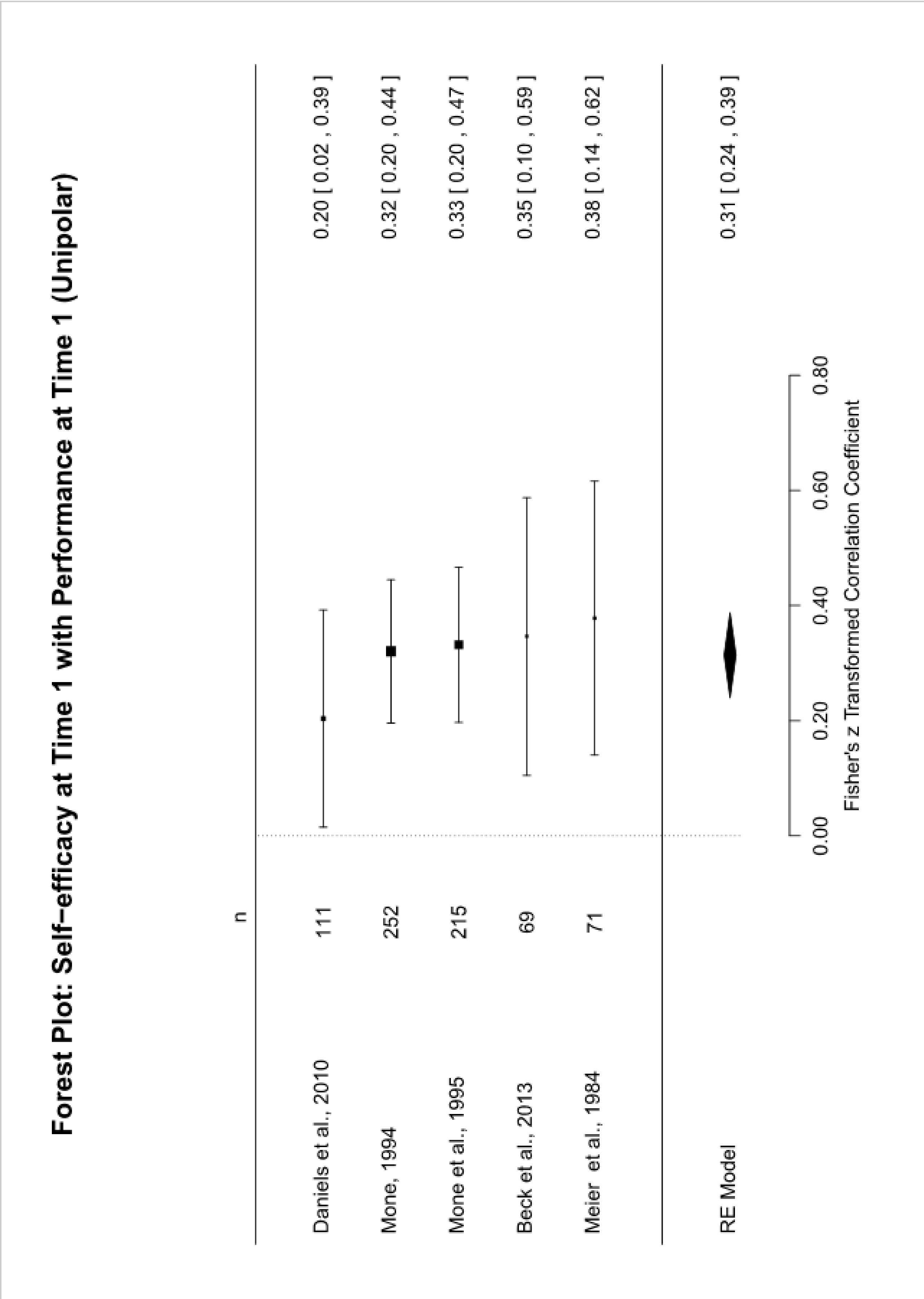
Funnel Plot: Self–efficacy at Time 1 with Performance at Time 2 (Unipolar)
Standard Error
0.000
0.031
0.062
0.092
0.123
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50



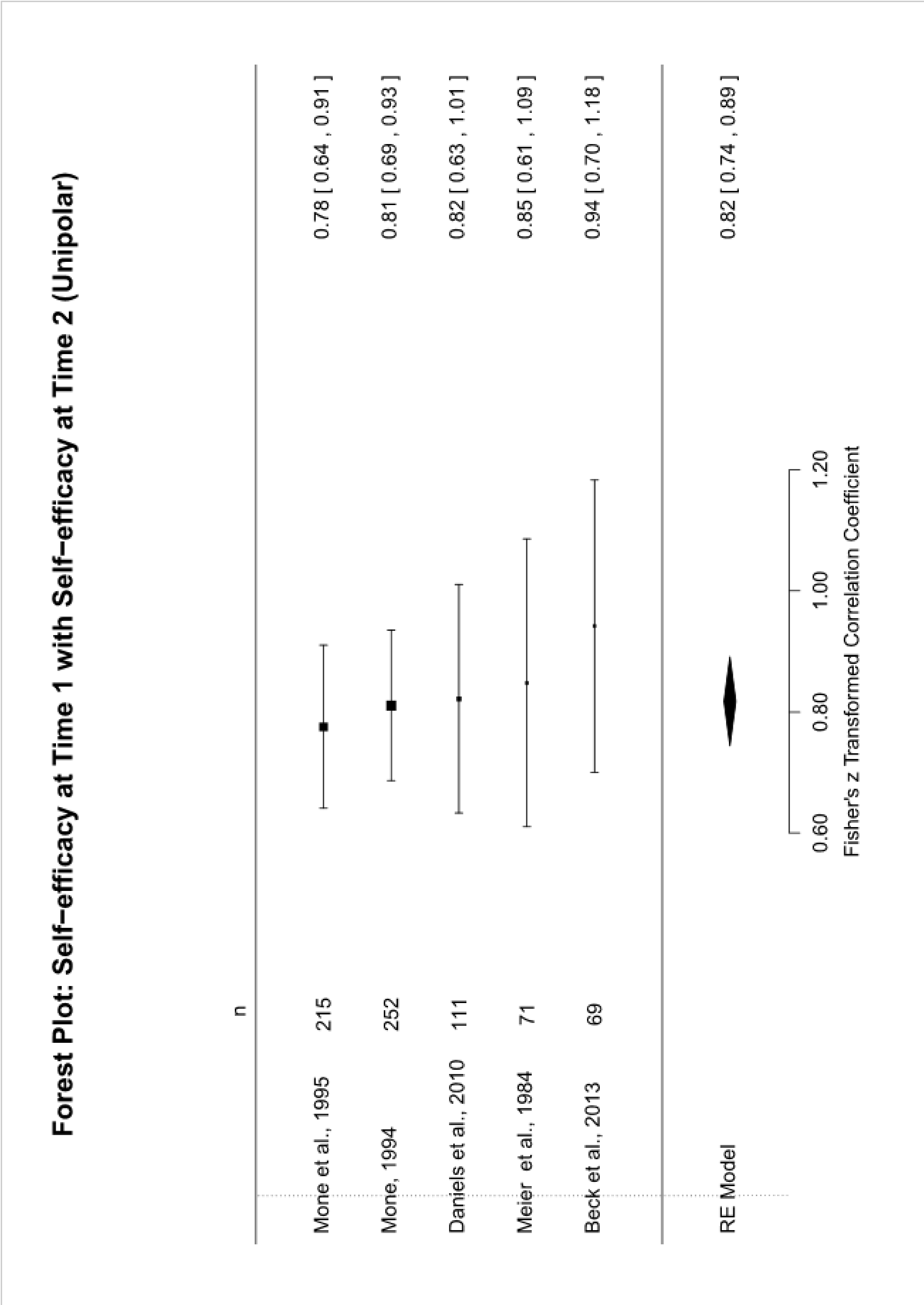
Funnel Plot: Self−efficacy at Time 2 with Performance at Time 1 (Unipolar)
Standard Error
0.000
0.031
0.062
0.092
0.123
Fisher's z Transformed Correlation Coefficient
0.20
0.30
0.40
0.50
0.60
0.70


Funnel Plot: Self−efficacy at Time 2 with Performance at Time 2 (Unipolar)
Standard Error
0.000
0.031
0.062
0.092
0.123
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50

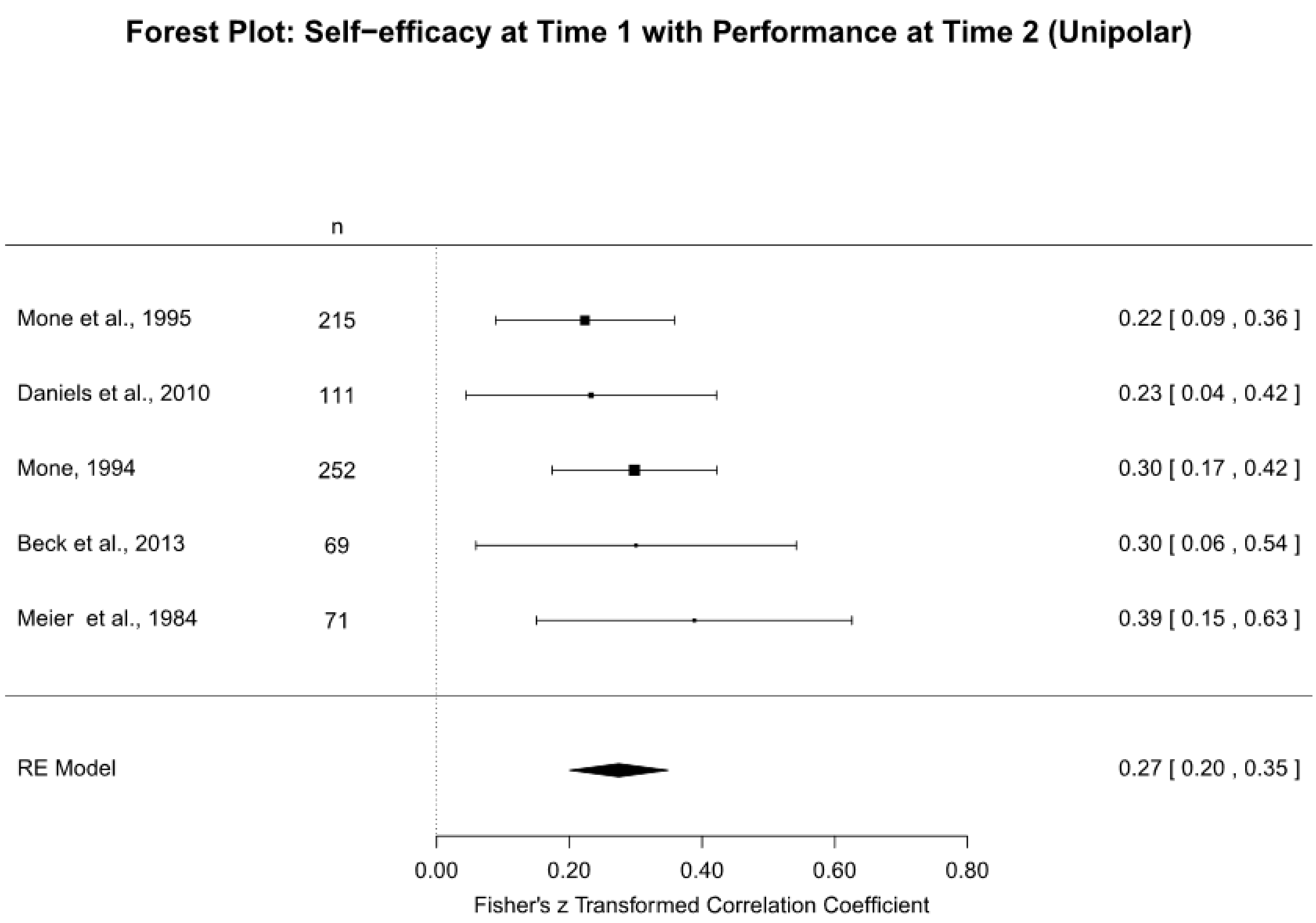

Funnel Plot: Performance at Time 1 with Performance at Time 2 (Unipolar)
Standard Error
0.000
0.031
0.062
0.092
0.123
0.30
0.40
0.50
0.60
0.70
0.80
0.90
Fisher's z Transformed Correlation Coefficient

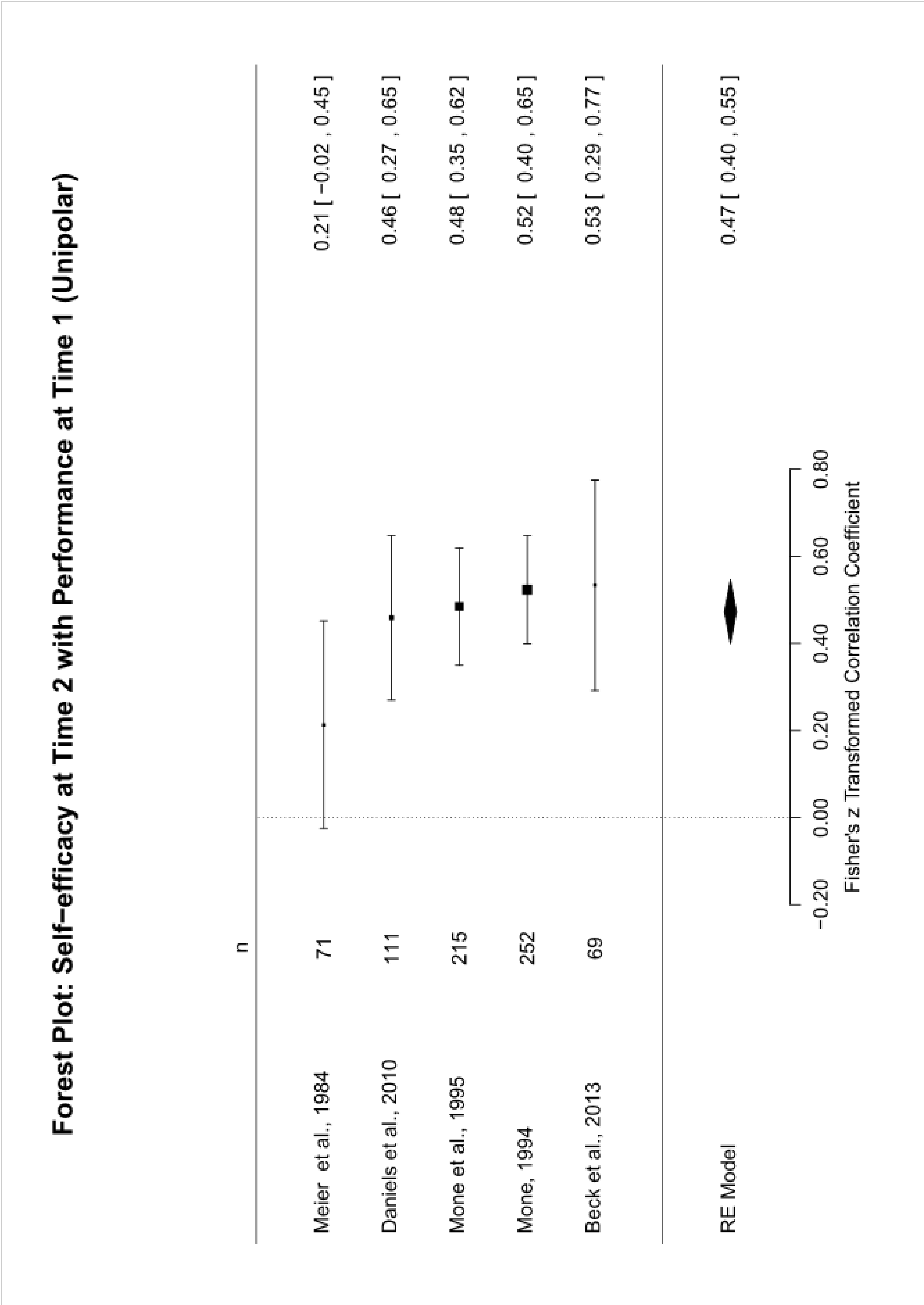


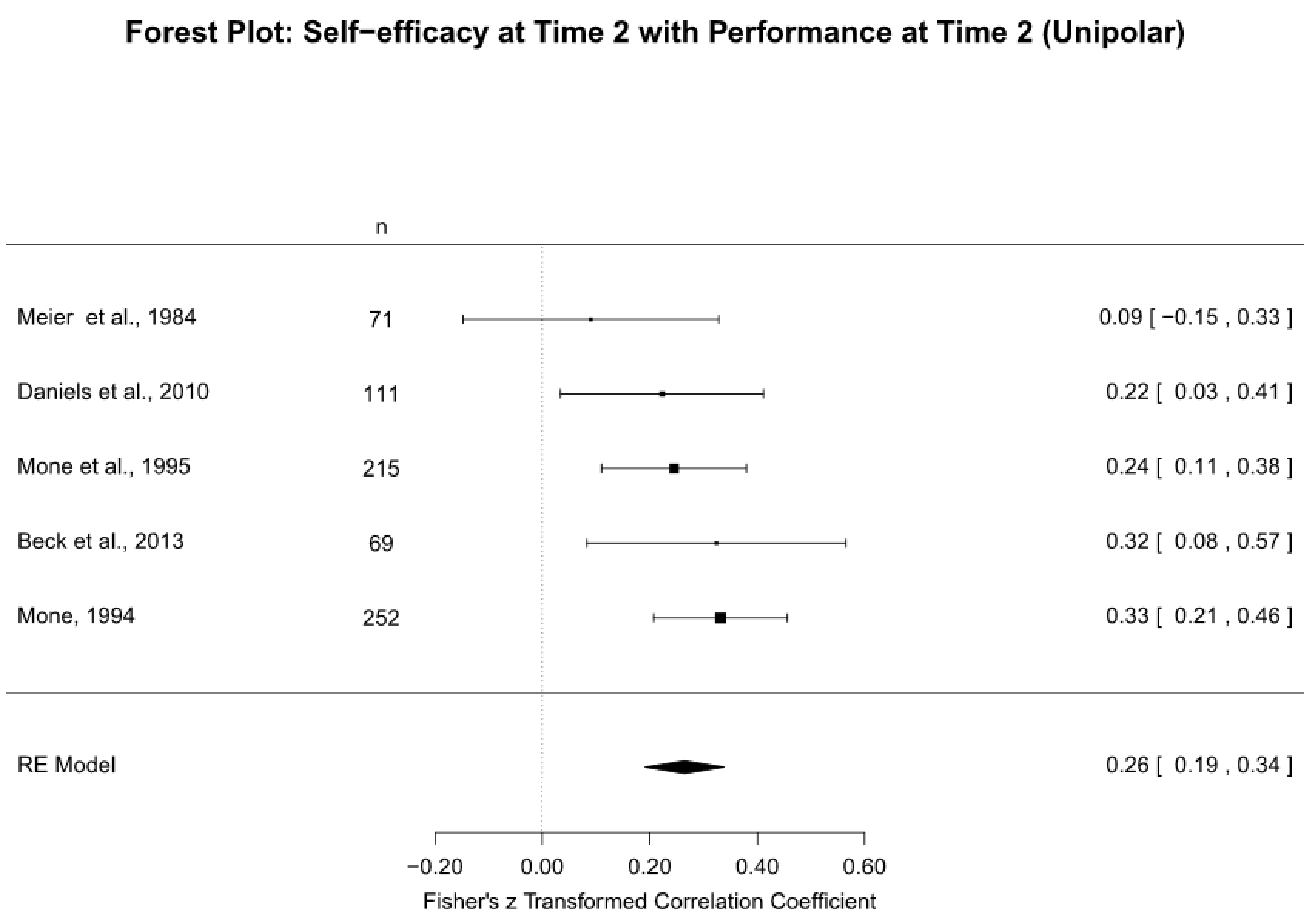
Forest Plot: Self–efficacy at Time 1 with Performance at Time 1 (Unipolar)
n
Daniels et al., 2010
111
0.20 [0.02 , 0.39]
Mone, 1994
252
0.32 [0.20 , 0.44]
Mone et al., 1995
215
0.33 [0.20 , 0.47]
Beck et al., 2013
69
0.35 [0.10 , 0.59]
Meier et al., 1984
71
0.38 [0.14 , 0.62]
RE Model
0.31 [0.24 , 0.39]
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient

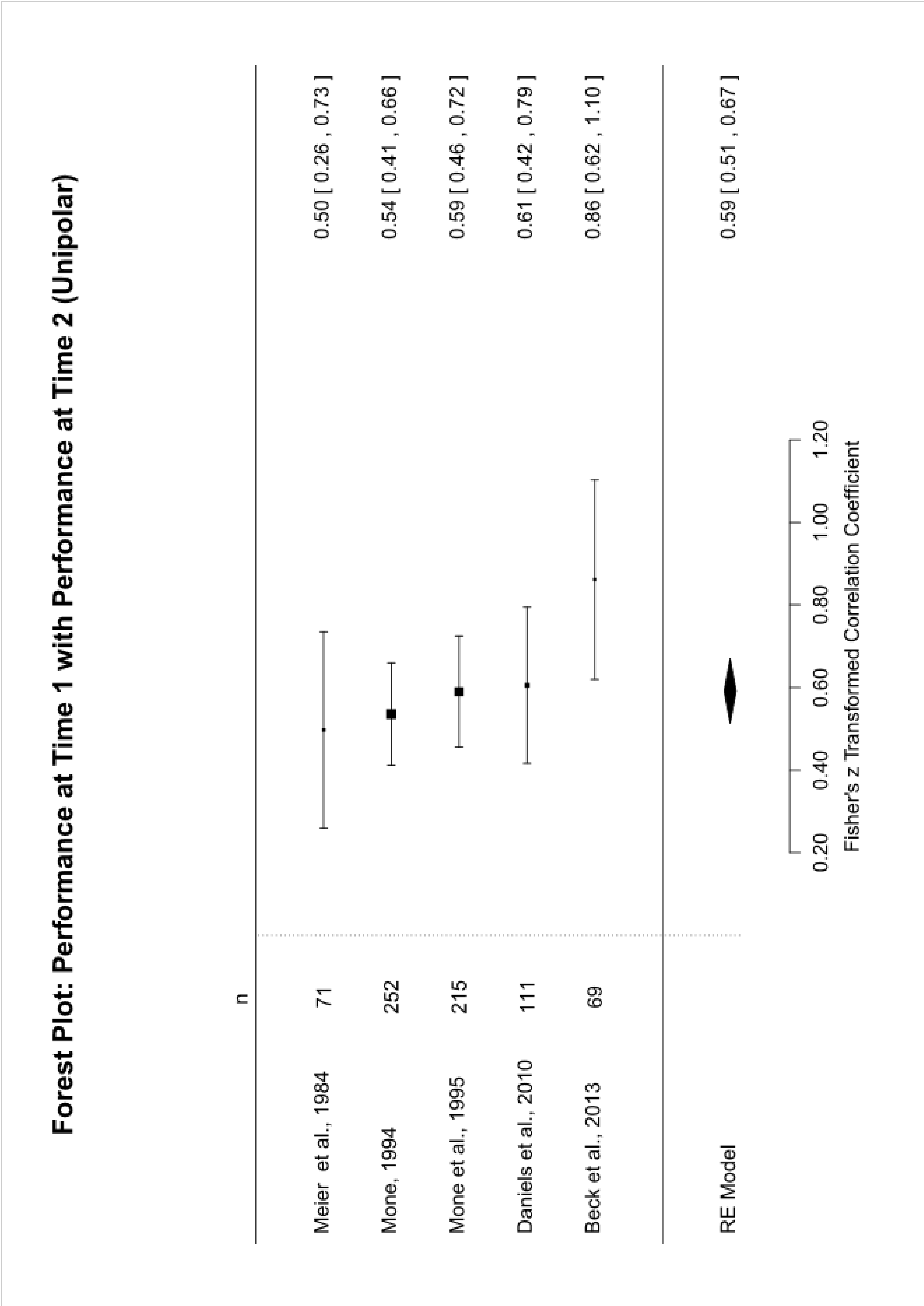


Forest Plot: Self–efficacy at Time 1 with Self–efficacy at Time 2 (Unipolar)
n
Mone et al., 1995 215 0.78 [0.64 , 0.91]
Mone, 1994 252 0.81 [0.69 , 0.93]
Daniels et al., 2010 111 0.82 [0.63 , 1.01]
Meier et al., 1984 71 0.85 [0.61 , 1.09]
Beck et al., 2013 69 0.94 [0.70 , 1.18]
RE Model 0.82 [0.74 , 0.89]
0.60
0.80
1.00
1.20
Fisher's z Transformed Correlation Coefficient


Forest Plot: Self−efficacy at Time 1 with Performance at Time 2 (Unipolar)
n
Mone et al., 1995
215
0.22 [0.09 , 0.36]
Daniels et al., 2010
111
0.23 [0.04 , 0.42]
Mone, 1994
252
0.30 [0.17 , 0.42]
Beck et al., 2013
69
0.30 [0.06 , 0.54]
Meier et al., 1984
71
0.39 [0.15 , 0.63]
RE Model
0.27 [0.20 , 0.35]
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient


Forest Plot: Self–efficacy at Time 2 with Performance at Time 1 (Unipolar)
n
Meier et al., 1984
71
0.21 [−0.02 , 0.45]
Daniels et al., 2010
111
0.46 [0.27 , 0.65]
Mone et al., 1995
215
0.48 [0.35 , 0.62]
Mone, 1994
252
0.52 [0.40 , 0.65]
Beck et al., 2013
69
0.53 [0.29 , 0.77]
RE Model
0.47 [0.40 , 0.55]
−0.20
0.00
0.20
0.40
0.60
0.80
Fisher's z Transformed Correlation Coefficient


Forest Plot: Self−efficacy at Time 2 with Performance at Time 2 (Unipolar)
n
Meier et al., 1984
71
0.09 [−0.15 , 0.33]
Daniels et al., 2010
111
0.22 [0.03 , 0.41]
Mone et al., 1995
215
0.24 [0.11 , 0.38]
Beck et al., 2013
69
0.32 [0.08 , 0.57]
Mone, 1994
252
0.33 [0.21 , 0.46]
RE Model
0.26 [0.19 , 0.34]
−0.20
0.00
0.20
0.40
0.60
Fisher's z Transformed Correlation Coefficient

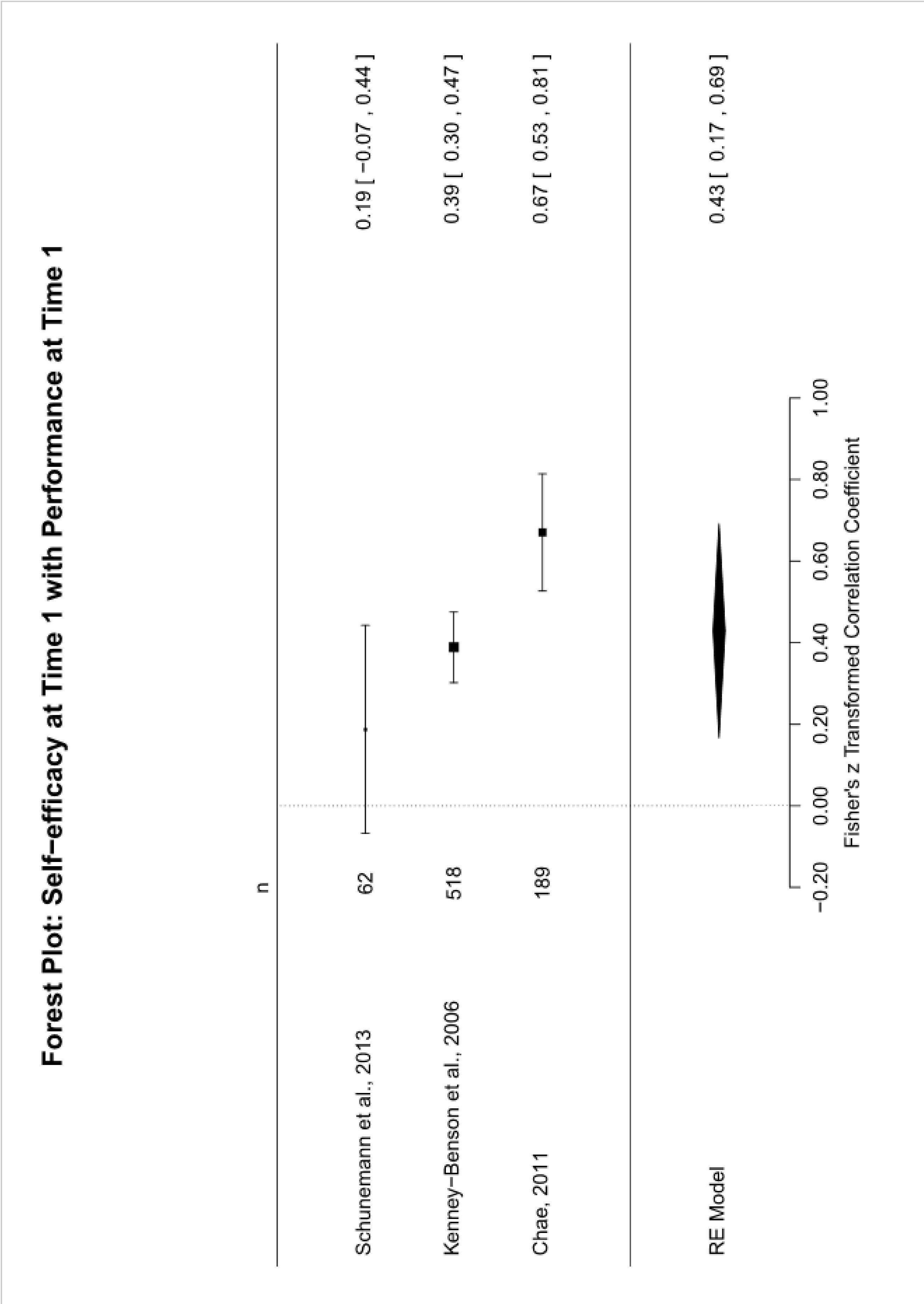

Forest Plot: Performance at Time 1 with Performance at Time 2 (Unipolar)
n
Meier et al., 1984
71
0.50 [0.26 , 0.73]
Mone, 1994
252
0.54 [0.41 , 0.66]
Mone et al., 1995
215
0.59 [0.46 , 0.72]
Daniels et al., 2010
111
0.61 [0.42 , 0.79]
Beck et al., 2013
69
0.86 [0.62 , 1.10]
RE Model
0.59 [0.51 , 0.67]
0.20
0.40
0.60
0.80
1.00
1.20
Fisher's z Transformed Correlation Coefficient

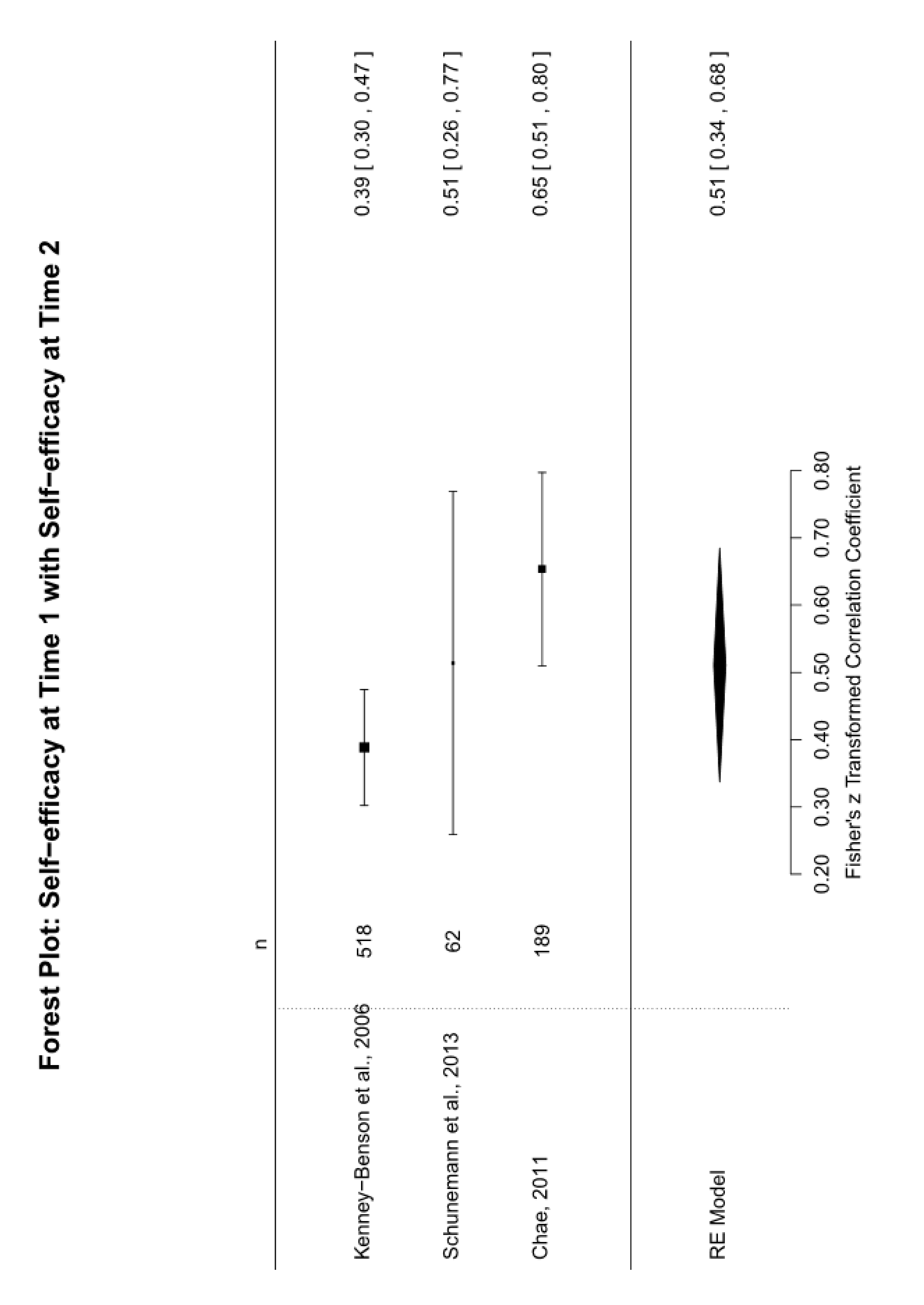
Appendix 2.8 **Model fit indices when cross-lagged paths constrained to be equal**

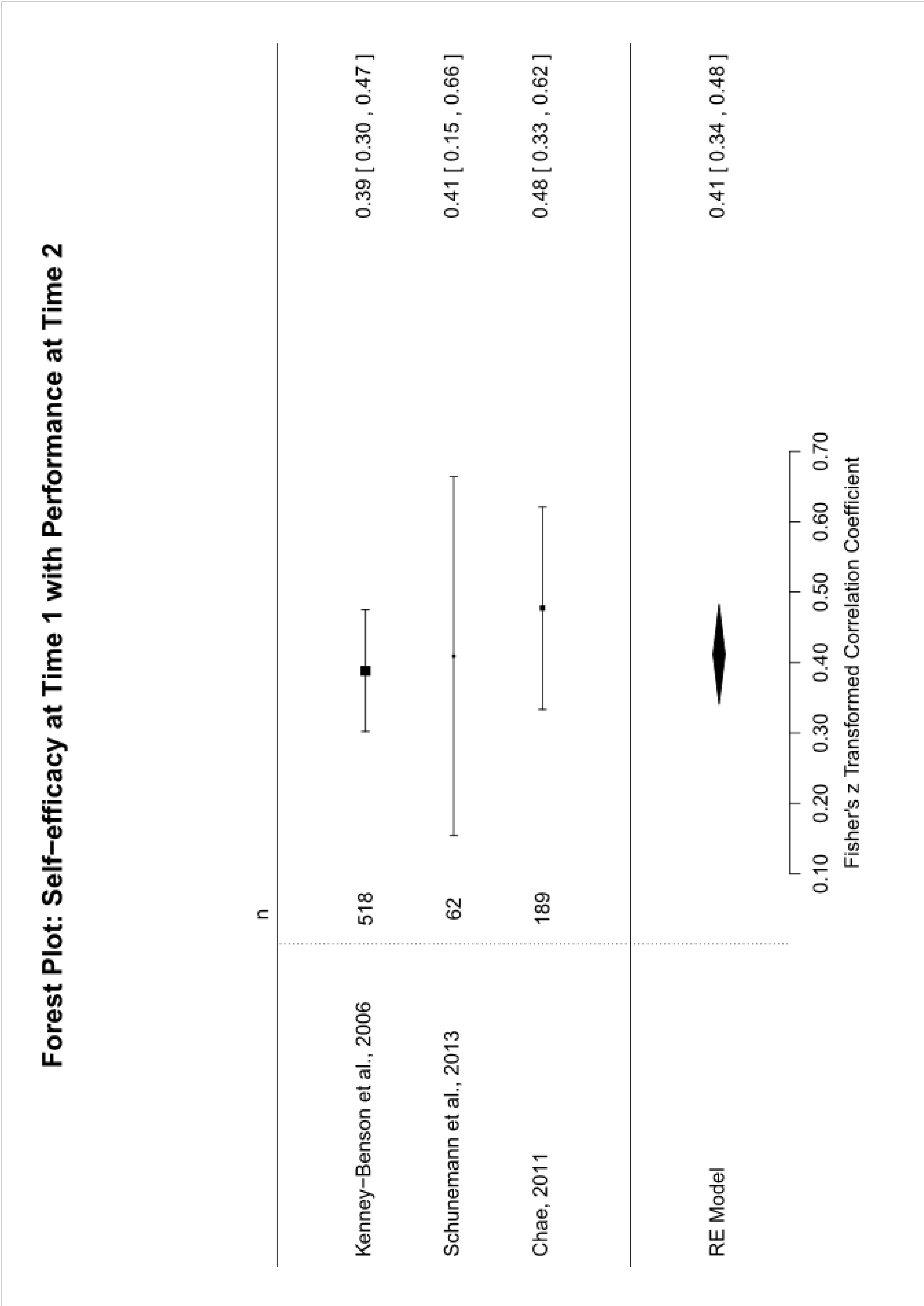
		n	χ^2	p
Age group	Adults	1182	20.586	<.0001
	Children	1506	17.56	<.0001
Lag	Short	1008	33.82	<.0001
	Long	1680	3.291	=.07
Match	Yes	969	12.058	=.0005
	No	1719	47.64	<.0001
Scale	Unipolar	718	14.69	<.0001
	Likert	1970	21.59	<.0001

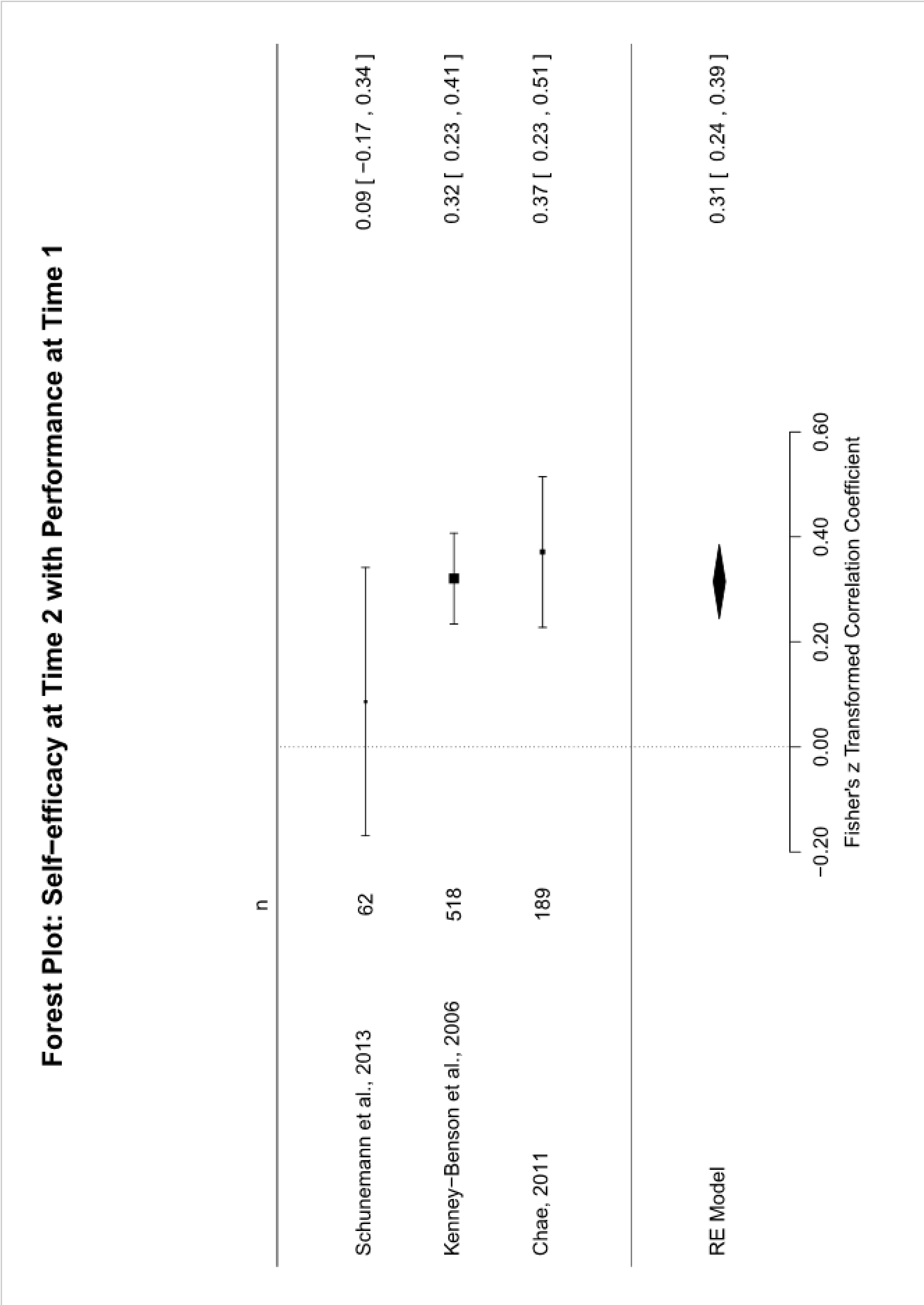
Note: df=1; Significant χ^2 test = a model with equal cross-lagged paths provided a significantly poorer fit to the data than the saturated model

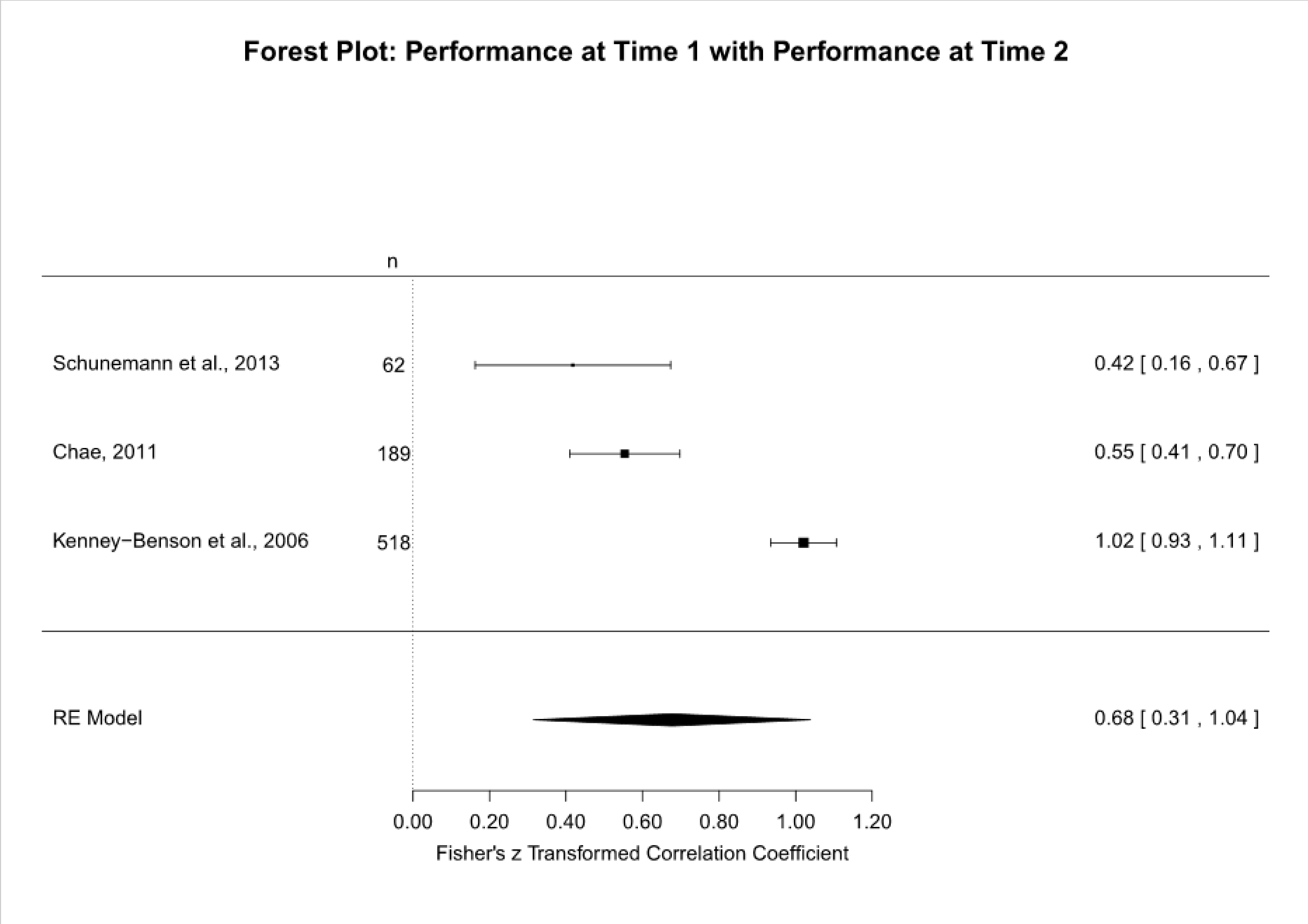
Appendix 3.1 Forest plots: Performance-first studies

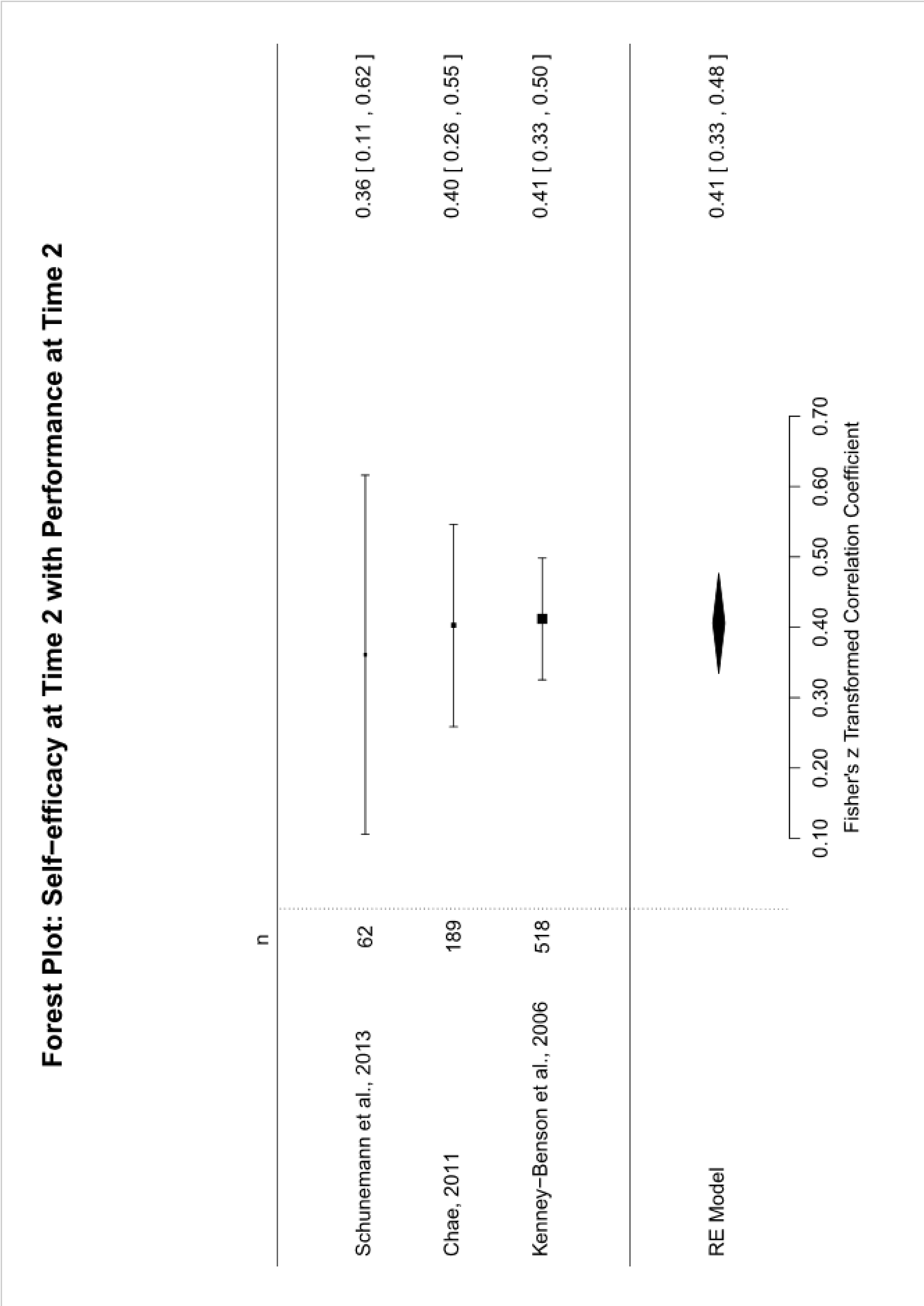

Forest Plot: Self–efficacy at Time 1 with Performance at Time 1
n
Schunemann et al., 2013
62
0.19 [−0.07 , 0.44]
Kenney−Benson et al., 2006
518
0.39 [0.30 , 0.47]
Chae, 2011
189
0.67 [0.53 , 0.81]
RE Model
0.43 [0.17 , 0.69]
−0.20
0.00
0.20
0.40
0.60
0.80
1.00
Fisher's z Transformed Correlation Coefficient


Forest Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2
n
Kenney−Benson et al., 2006
518
0.39 [0.30 , 0.47]
Schunemann et al., 2013
62
0.51 [0.26 , 0.77]
Chae, 2011
189
0.65 [0.51 , 0.80]
RE Model
0.51 [0.34 , 0.68]
0.20
0.30
0.40
0.50
0.60
0.70
0.80
Fisher's z Transformed Correlation Coefficient

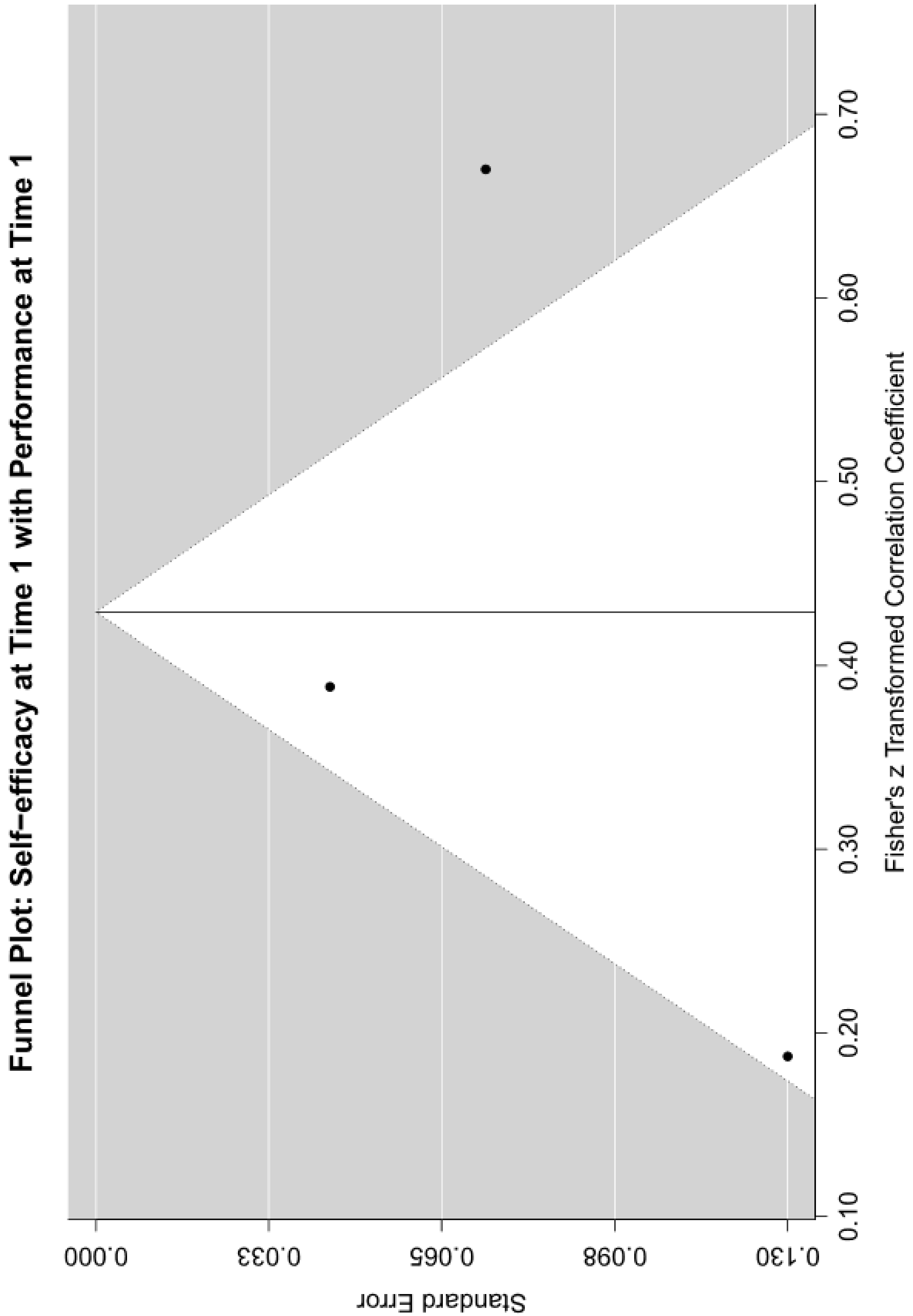

Forest Plot: Self−efficacy at Time 1 with Performance at Time 2
n
Kenney−Benson et al., 2006
518
0.39 [0.30 , 0.47]
Schunemann et al., 2013
62
0.41 [0.15 , 0.66]
Chae, 2011
189
0.48 [0.33 , 0.62]
RE Model
0.41 [0.34 , 0.48]
0.10
0.20
0.30
0.40
0.50
0.60
0.70
Fisher's z Transformed Correlation Coefficient

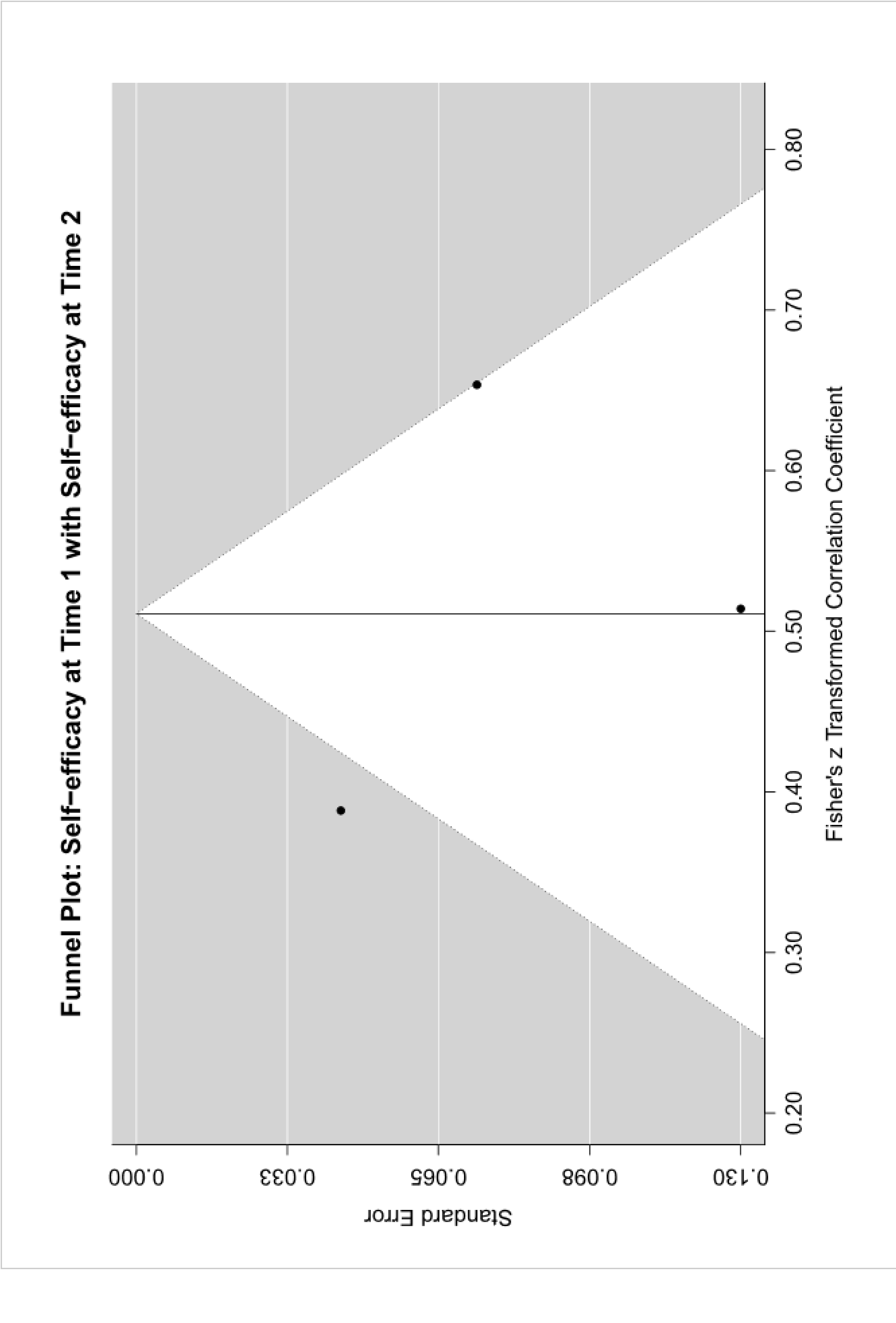

Forest Plot: Self−efficacy at Time 2 with Performance at Time 1
n
Schunemann et al., 2013
62
0.09 [−0.17 , 0.34]
Kenney−Benson et al., 2006
518
0.32 [0.23 , 0.41]
Chae, 2011
189
0.37 [0.23 , 0.51]
RE Model
0.31 [0.24 , 0.39]
−0.20
0.00
0.20
0.40
0.60
Fisher's z Transformed Correlation Coefficient

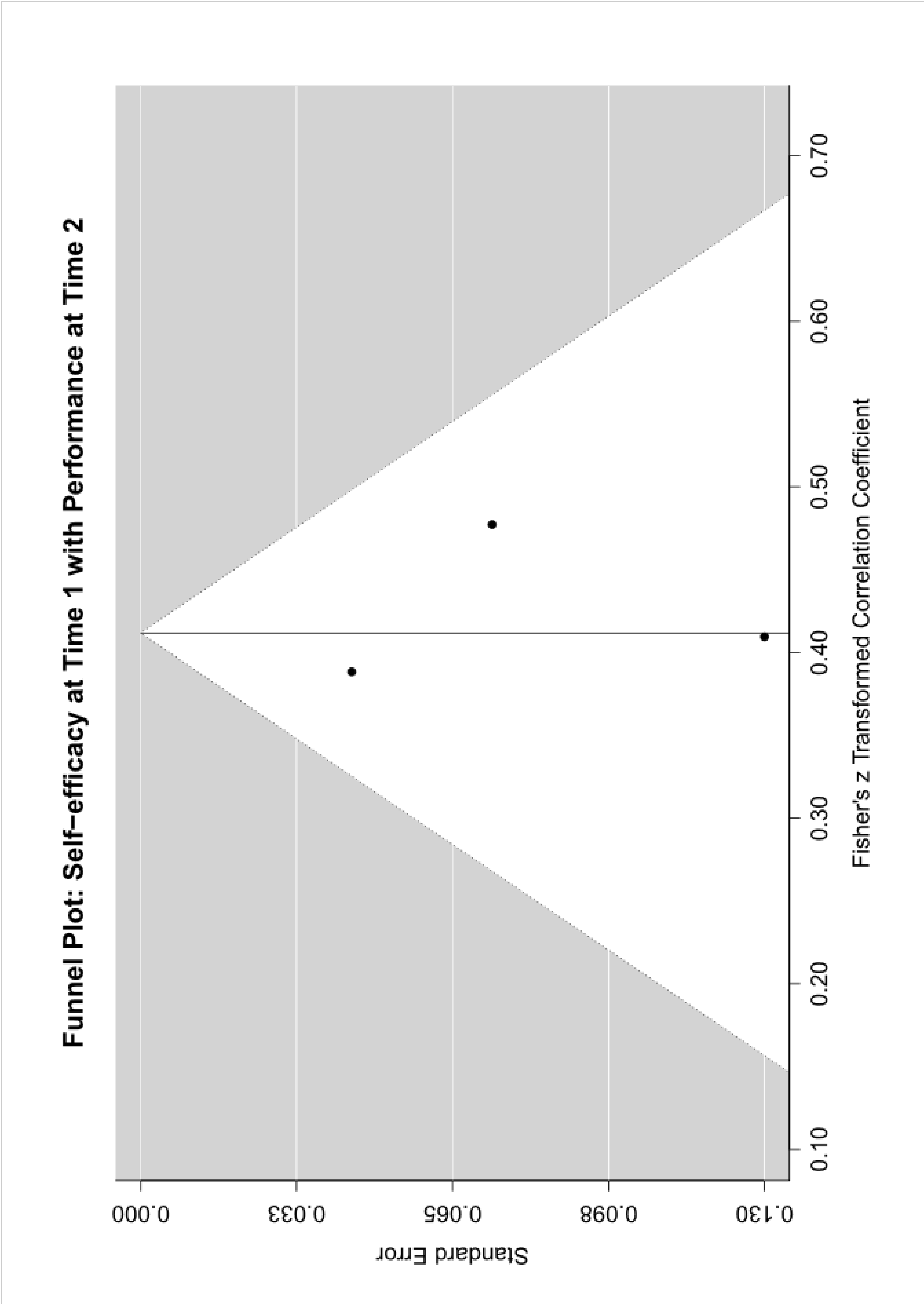

Forest Plot: Performance at Time 1 with Performance at Time 2
n
Schunemann et al., 2013
62
0.42 [0.16 , 0.67]
Chae, 2011
189
0.55 [0.41 , 0.70]
Kenney−Benson et al., 2006
518
1.02 [0.93 , 1.11]
RE Model
0.68 [0.31 , 1.04]
0.00
0.20
0.40
0.60
0.80
1.00
1.20
Fisher's z Transformed Correlation Coefficient

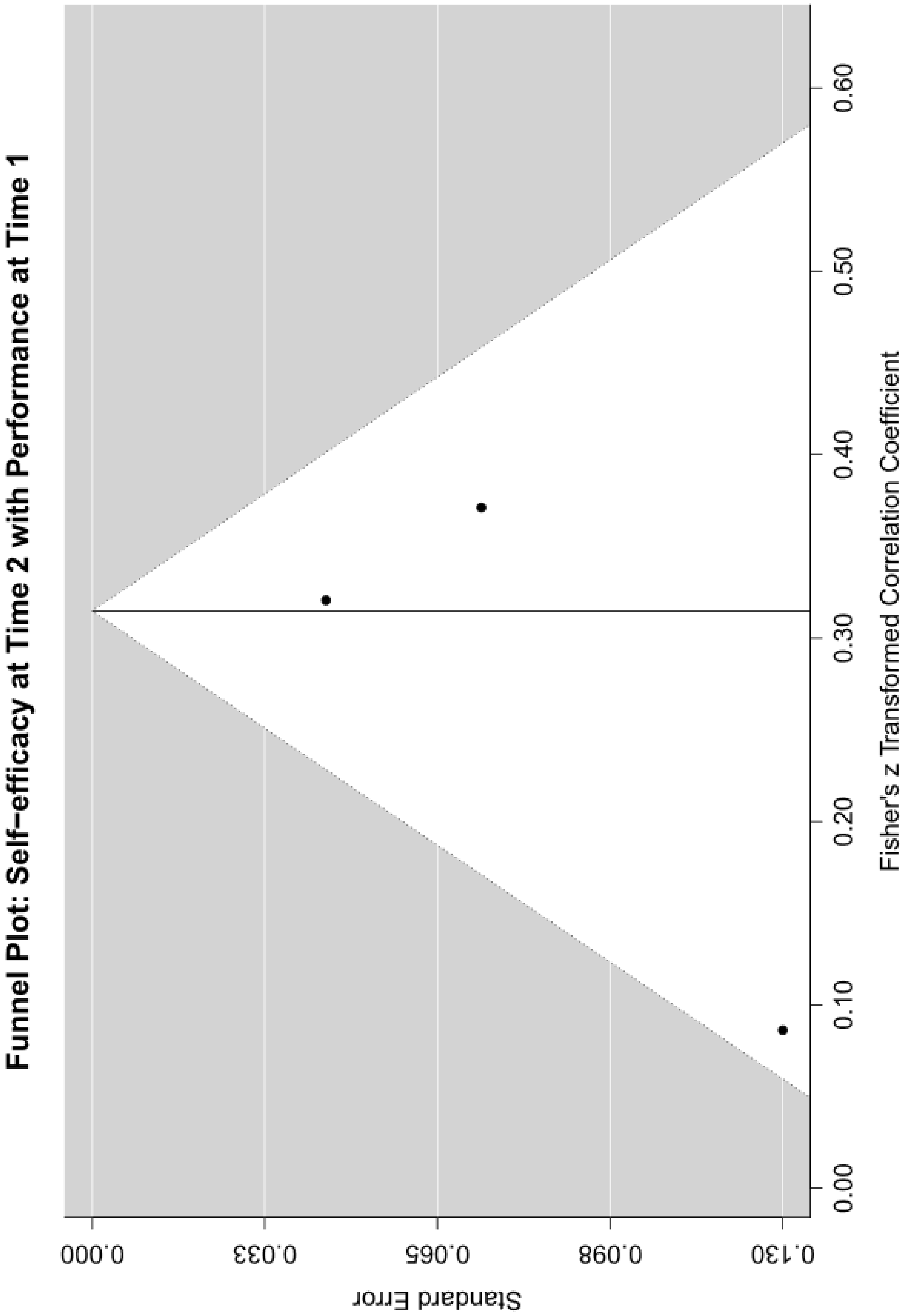

Forest Plot: Self−efficacy at Time 2 with Performance at Time 2
n
Schunemann et al., 2013
62
0.36 [0.11 , 0.62]
Chae, 2011
189
0.40 [0.26 , 0.55]
Kenney−Benson et al., 2006
518
0.41 [0.33 , 0.50]
RE Model
0.41 [0.33 , 0.48]
0.10
0.20
0.30
0.40
0.50
0.60
0.70
Fisher's z Transformed Correlation Coefficient

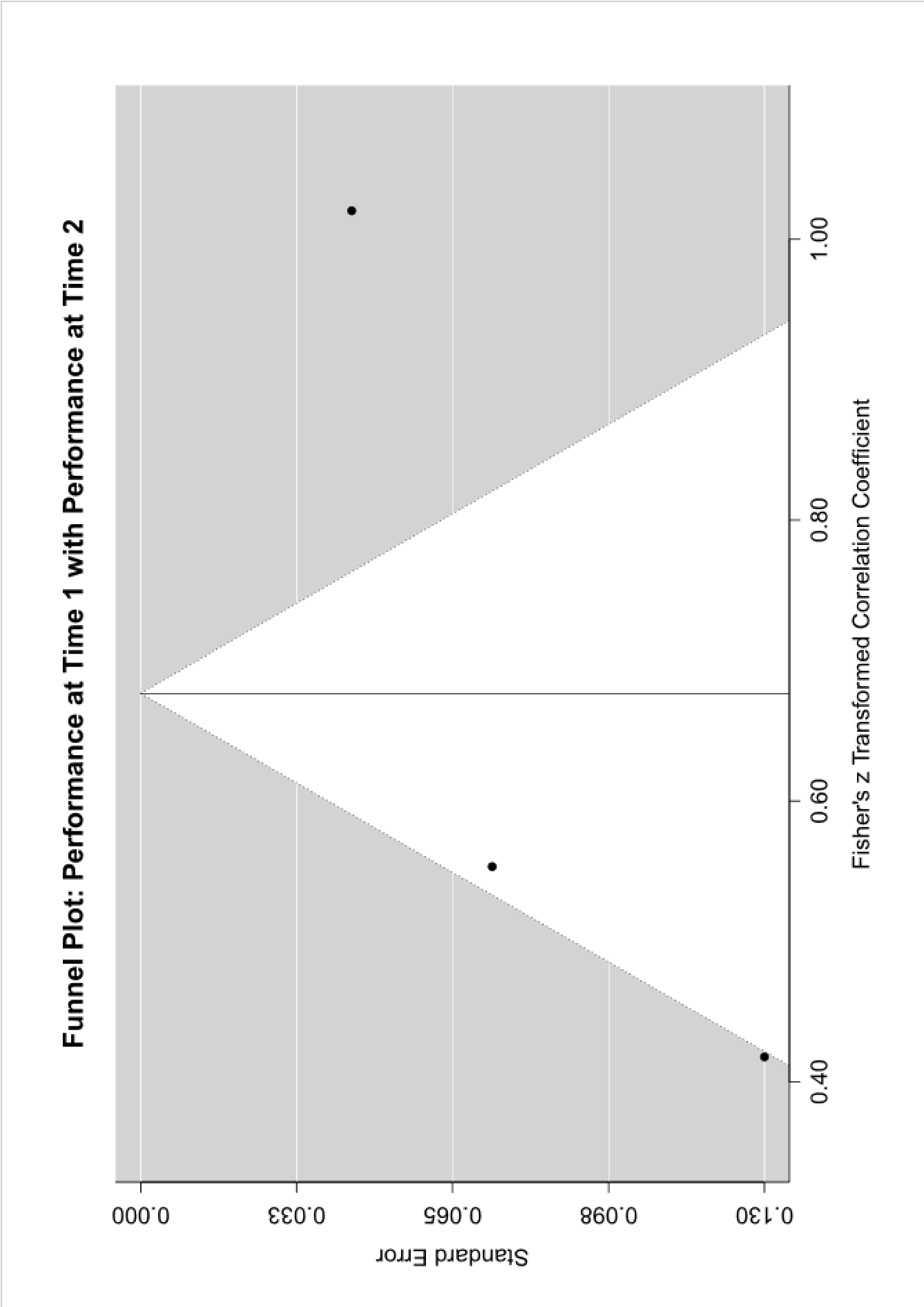
Appendix 3.2 Funnel plots: Performance-first studies

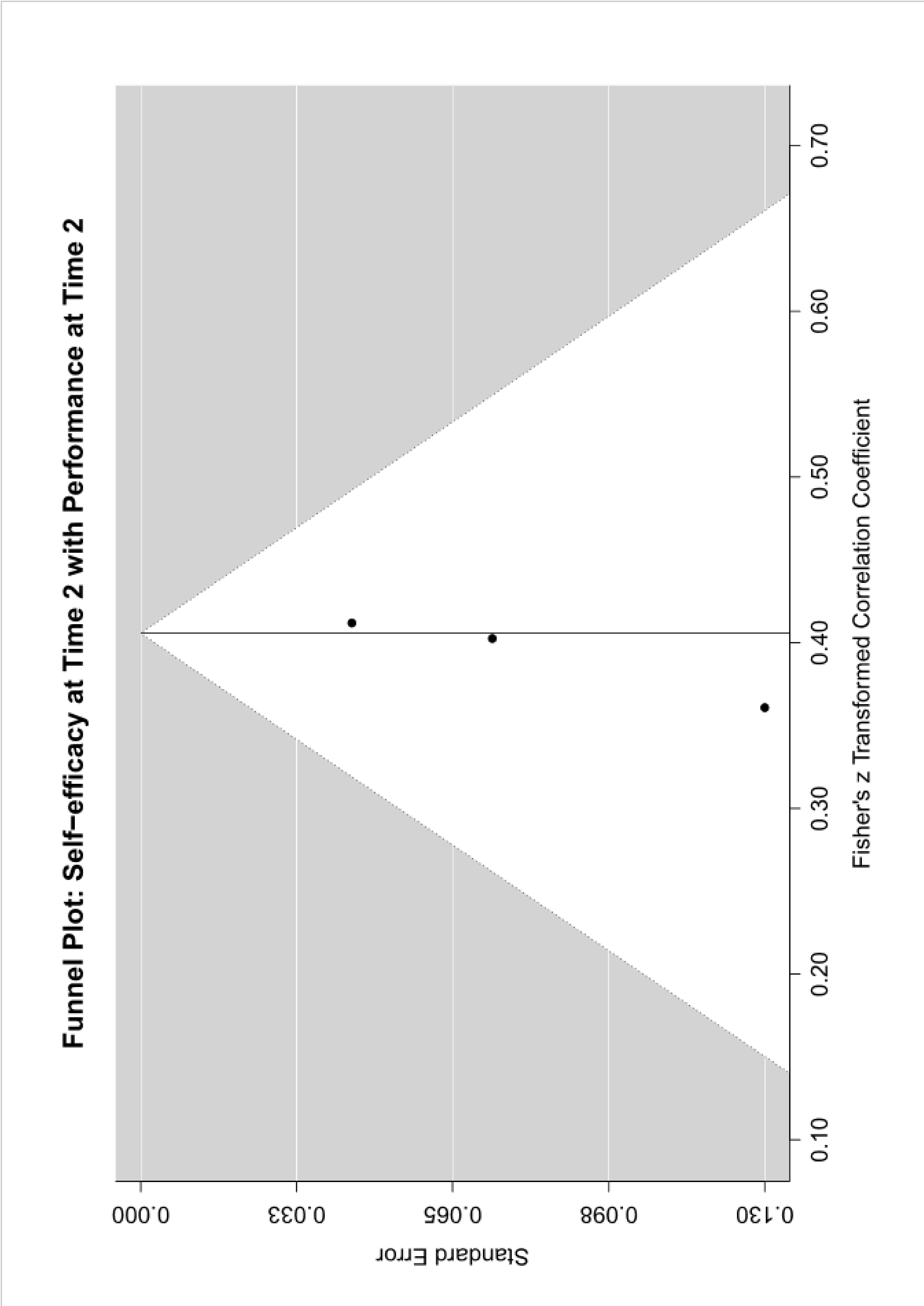

Funnel Plot: Self−efficacy at Time 1 with Performance at Time 1
Standard Error
0.000
0.033
0.065
0.098
0.130
Fisher's z Transformed Correlation Coefficient
0.10
0.20
0.30
0.40
0.50
0.60
0.70


Funnel Plot: Self−efficacy at Time 1 with Self−efficacy at Time 2
Standard Error
0.000
0.033
0.065
0.098
0.130
Fisher's z Transformed Correlation Coefficient
0.20
0.30
0.40
0.50
0.60
0.70
0.80


Funnel Plot: Self−efficacy at Time 1 with Performance at Time 2
Standard Error
0.000
0.033
0.065
0.098
0.130
0.10
0.20
0.30
0.40
0.50
0.60
0.70
Fisher's z Transformed Correlation Coefficient


Funnel Plot: Self−efficacy at Time 2 with Performance at Time 1
Standard Error
0.000
0.033
0.065
0.098
0.130
Fisher's z Transformed Correlation Coefficient
0.00
0.10
0.20
0.30
0.40
0.50
0.60


Funnel Plot: Performance at Time 1 with Performance at Time 2
Standard Error
0.000
0.033
0.065
0.098
0.130
Fisher's z Transformed Correlation Coefficient
0.40
0.60
0.80
1.00


Funnel Plot: Self−efficacy at Time 2 with Performance at Time 2
Standard Error
0.000
0.033
0.065
0.098
0.130
Fisher's z Transformed Correlation Coefficient
0.10
0.20
0.30
0.40
0.50
0.60
0.70